

# Contrast summation across eyes and space is revealed along the entire dipper function by a “Swiss cheese” stimulus

Tim S. Meese

School of Life and Health Sciences, Aston University,  
Birmingham, UK



Daniel H. Baker

School of Life and Health Sciences, Aston University,  
Birmingham, UK



Previous contrast discrimination experiments have shown that luminance contrast is summed across ocular (T. S. Meese, M. A. Georgeson, & D. H. Baker, 2006) and spatial (T. S. Meese & R. J. Summers, 2007) dimensions at threshold and above. However, is this process sufficiently general to operate across the *conjunction* of eyes and space? Here we used a “Swiss cheese” stimulus where the blurred “holes” in sine-wave carriers were of equal area to the blurred target (“cheese”) regions. The locations of the target regions in the monocular image pairs were interdigitated across eyes such that their binocular sum was a uniform grating. When pedestal contrasts were above threshold, the monocular neural images contained strong evidence that the high-contrast regions in the two eyes did not overlap. Nevertheless, sensitivity to dual contrast increments (i.e., to contrast increments in different locations in the two eyes) was a factor of  $\sim 1.7$  greater than to single increments (i.e., increments in a single eye), comparable with conventional binocular summation. This provides evidence for a contiguous area summation process that operates at all contrasts and is influenced little, if at all, by eye of origin. A three-stage model of contrast gain control fitted the results and possessed the properties of ocularity invariance and area invariance owing to its cascade of normalization stages. The implications for a population code for pattern size are discussed.

**Keywords:** human vision, contrast discrimination, masking, contrast gain control, interocular suppression, rivalry, surround suppression

**Citation:** Meese, T. S., & Baker, D. H. (2011). Contrast summation across eyes and space is revealed along the entire dipper function by a “Swiss cheese” stimulus. *Journal of Vision*, 11(1):23, 1–23, <http://www.journalofvision.org/content/11/1/23>, doi:10.1167/11.1.23.

## Introduction

An important property of mammalian vision is neuronal convergence. As one moves up the visual hierarchy, visual mechanisms are imbued with increasingly elaborate properties owing to their feeds from arrays of lower order mechanisms, each selective for different regions along the dimension(s) of interest. Physiologists and anatomists see this at all levels of the visual system, including the retina (photoreceptors to retinal ganglion cells), the projection to V1 (LGN cells to simple cells), within V1 itself (simple cells to complex cells), intermediate areas such as MT (velocity tuning and pattern cells), V4 (spatial form), IT/LO (texture and objects), and so forth.

For those who study human spatial vision using contrast detection and contrast discrimination tasks, the brain’s ubiquitous neuronal convergence poses two puzzles. First, why is the benefit of increasing the area of a sine-wave grating so slight at detection threshold (Robson & Graham, 1981)? Second, why is that benefit completely abolished when the task is raised above threshold,

becoming one of contrast discrimination (Legge & Foley, 1980)?

## Orthodoxy

Early models of spatial vision implied that in spite of neuronal convergence elsewhere in the system, there is no benefit from neuronal convergence beyond V1 receptive field size for contrast detection and discrimination tasks (Legge & Foley, 1980). The standard model proposed that beyond the size of receptive fields in V1 ( $\sim 2$  cycles of sine-wave grating), the benefit of grating area at detection threshold is merely probabilistic: the greater the number of independent mechanisms excited by the stimulus, the greater the probability of detection. This interpretation of spatial summation has been influential, but there is no direct psychophysical evidence to support it; only the circumstantial evidence that the relevant data are consistent with the levels of summation predicted by probability summation (Meese & Williams, 2000; Robson & Graham, 1981). We refer to this as the *first dogma of spatial vision*.

An important study that extended the investigation of spatial vision above threshold was conducted by Legge and Foley (1980). They compared contrast discrimination thresholds for small and large grating areas but found no difference once the pedestal contrast was well above threshold. They suggested that the benefit of area is lost above detection threshold because the noise among the detecting mechanisms becomes correlated, thereby nullifying the statistical advantage of the selection process (e.g., a MAX operator). However, Legge and Foley (1980) manipulated the area of both the target and the pedestal, leaving open the possibility that the process of area summation remains operative for the target but is nullified by extra suppression from the pedestal, which also increases in size (Bonneh & Sagi, 1999; Meese, Hess, & Williams, 2005). Therefore, we refer to the belief that there is no area summation of signal contrast well above detection threshold (Cannon & Fullenkamp, 1988; Legge & Foley, 1980; McIlhagga & Pääkkönen, 1999; Swanson, Wilson, & Giese, 1984) as the *second dogma of spatial vision*.

## Probability summation, signal selection, and the MAX operator

The term “probability summation” is used widely in the spatial vision literature, though here we offer an alternative term: “signal selection,” which has greater neutrality regarding the neural underpinning. For example, when a signal line wins out probabilistically (by responding above “threshold”), it is *self-selecting* from the competing signal lines. Similarly, we conceptualize the MAX operation as the system *selecting* one of several signal lines (channels), either for decision making (Pelli, 1985) or further transmission (Riesenhuber & Poggio, 1999).

The MAX operator is of particular interest in neuroscience (Finn & Ferster, 2007; Lampl, Ferster, Poggio, & Riesenhuber, 2004) because it provides a method by which various stimulus invariances can be achieved (Riesenhuber & Poggio, 1999, 2002). In conjunction with stimulus selectivity from linear summation (Ghose & Maunsell, 2008), this has been central to models of object and shape recognition in V4 and higher levels of the visual hierarchy (Cadieu et al., 2007; Kouh & Poggio, 2008). The MAX operator has also attracted interest in psychophysics. For example, when it is performed over independently noisy input lines (Pelli, 1985), the process has properties similar to the probability summation model of Sachs, Nachmias, and Robson (1971) and the vector magnitude model of Quick (1974). Thus, the MAX operator is a viable implementation of signal selection and represents a contemporary treatment of probability summation (Pelli, 1985; Tyler & Chen, 2000), even though it does not involve *summing* probabilities directly (Tyler & Chen, 2000). Nevertheless, owing to the widespread use of the term “probability summation,” we shall treat it

interchangeably with the term “signal selection,” in this paper at least.

## Challenges to orthodoxy

In spite of the success of the early models of spatial vision and probability summation (Bergen, Wilson, & Cowen, 1979; Legge & Foley, 1980; Robson & Graham, 1981; Sachs et al., 1971; Wilson, McFarlane, & Phillips, 1983), challenges have been raised to the orthodox positions on both area (spatial) probability summation and the loss of area summation above threshold (the first and second dogmas of spatial vision). We outline these challenges in the next two subsections.

### Signal combination for contrast at threshold

Probability summation (signal selection) models are often approximated using a Minkowski metric with a Minkowski exponent of about 4 (Robson & Graham, 1981; Tyler & Chen, 2000; Watson, 1979). However, care is needed with this implementation (Pelli, 1985; Tyler & Chen, 2000). For example, Meese and Summers (2009) showed that when placed in a spatial pooling model it failed (badly) to predict the combined results of the level of summation *and* the slope of the psychometric function. As eluded earlier, a further problem is that although probability summation is consistent with the weak levels of summation that it is proposed to explain (Robson & Graham, 1981) other—very different—processes also predict similar levels of summation. For example, Rovamo et al. presented a series of studies that proposed spatial integration by a matched filter (Luntinen, Rovamo, & Näsänen, 1995; Näsänen, Tiippana, & Rovamo, 1998; Rovamo, Luntinen, & Nasanen, 1993; Rovamo, Mustonen, & Nasanen, 1994; Rovamo, Ukkonen, Thompson, & Nasanen, 1994; see also Manahilov, Simpson, & McCulloch, 2001). The fits to their data were very good, though their analysis did not displace the orthodox account. In a more recent study, Foley, Varadharajan, Koh, and Farias (2007) performed area summation experiments and a detailed mathematical analysis and concluded that multiple receptive fields are summed after each being subject to an accelerating contrast-response nonlinearity and retinal inhomogeneity. However, Minkowski summation remained a valid interpretation of their results. Meese and Summers (2007) proposed a similar arrangement to Foley<sup>1</sup> (see also Meese, 2004, 2010; Meese et al., 2005) but also provided supporting evidence from the slope of the psychometric function. Using this additional constraint (and the “Swiss cheese” stimuli described below), they showed that their area summation results were not explained by a MAX rule or Minkowski summation with exponents in the conventional “probability summation” range of 3 or 4. This was elaborated further by the detailed analysis and experiments of Meese (2010) and Meese and

Summers (2009), who proposed an overall transducer exponent of about 2 (around “threshold”) prior to summation of contrast over area. Finally, several studies have investigated contrast summation for strips of grating across parafoveal retinal regions, where sensitivity is more homogeneous than in the fovea (Manahilov et al., 2001; Mayer & Tyler, 1986; Meese & Hess, 2007). All of these studies found higher levels of summation than the earlier study of Robson and Graham (see Meese & Hess, 2007 for a review) suggesting that contrast detection involves a process more potent than signal selection, at least in some circumstances.

Overall then, current evidence dispels the first dogma of spatial vision and favors a strategy involving the linear sum of contrast over area, following an accelerating relation between signal contrast and performance (Meese, 2010; Meese & Summers, 2007, 2009). This form of summation is sometimes referred to as “physiological summation,” though we prefer the term “signal combination” as this makes it distinct from the alternative MAX rule (signal selection), which, in principle, could have a physiological implementation at an early sensory stage (Finn & Ferster, 2007; Lampl et al., 2004).

#### **Area summation of contrast operates at and above detection threshold**

Recent evidence has also dispelled the second dogma of spatial vision: area summation of contrast does take place above detection threshold after all. In addition to the threshold work described above, Meese and Summers (2007) measured dipper functions (contrast discrimination functions, or TvC curves) using a large pedestal (22.5 cycles of 2.5 c/deg grating). Their targets were contrast increments applied to either (i) the entire pedestal or (ii) multiple patches across the pedestal, which we refer to as a “Swiss cheese” stimulus (owing to the holes; see forward to Figure 2). With this arrangement, the sum of carrier contrast over area for the pedestal increments was twice that of the Swiss cheese increments. Crucially, Meese and Summers (2007) found that sensitivity to full pedestal increments (in terms of carrier contrast) was also nearly twice that of Swiss cheese increments, and that this happened along the entire dipper function. This provided the first clear evidence for an area summation process for luminance contrast that operates well above detection threshold in the central visual field (but see also Bonnef & Sagi, 1999; Levi & Klein, 2000; Meese, 2004; Meese et al., 2005; Wilkinson, Wilson, & Ellemberg, 1997). Furthermore, the results were predicted by a contrast gain control model involving linear spatial summation across multiple classical receptive fields and countersuppression from those same receptive fields (Meese & Summers, 2007; Wilkinson et al., 1997). Not only did this model account for the novel result described above, but also Legge and Foley’s (1980) orthodox result that dipper handles converge when the contrast area of both the target

and the pedestal are increased. (We consider this in more detail in Appendix B.)

#### **Binocular summation operates at and above detection threshold**

Another form of contrast summation is that which takes place between the eyes, usually referred to as binocular summation. Binocular summation of contrast at threshold can be substantial (a factor of 1.7; Baker, Meese, Mansouri, & Hess, 2007; Baker, Meese, & Summers, 2007; Georgeson & Meese, 2005; Meese, Challinor, & Summers, 2008; Meese, Georgeson, & Baker, 2006; Meese & Summers, 2009; Rose, 1980; Simmons, 2005). Certainly, it can be significantly greater than the factor of  $\sqrt{2}$  (3 dB) predicted by models of the ideal observer (Campbell & Green, 1965) and quadratic summation (Legge, 1984; see Meese & Summers, 2009) implying that performance-limiting noise is placed beyond binocular summation. In a binocular masking experiment analogous to the area summation experiment in the previous subsection, Meese et al. (2006) measured summation across eyes for patches of grating at and above detection threshold. Their pedestals were binocular patches of 1 c/deg grating (5 cycles) and target increments were presented to either one eye or both. For the binocular increment, they found that sensitivity was almost a factor of two greater than for the monocular increment over the entire dipper function. This provided good evidence that binocular summation of contrast takes place for the entire suprathreshold range, just like area summation.

#### **Preliminary outline of the current work**

A natural question arises from the studies above. Given that summation of contrast extends across area (Meese & Summers, 2007) and eyes (Meese et al., 2006), might it extend across the conjunction of area and eyes? That is, can visual signals from different areas in different eyes be combined? In fact, this has been confirmed around detection threshold by Meese and Summers (2009), but they did not consider pedestal masking. Here we extend the inquiry to suprathreshold vision by measuring contrast discrimination functions for a wide range of pedestal contrasts (i.e., we measured dipper functions). The logic of the stimulus design and its relation to our two previous studies with dipper functions is outlined in Figure 1. Note that within each of the three studies the pedestal is constant across conditions, only the configuration of the target varies. We consider this further in the Discussion section.

For the experiment here (Figure 1c), contrast increments were presented to one eye in one location, the other eye in the other location, or both of these. For convenience, we refer to the first two conditions as single increments, and



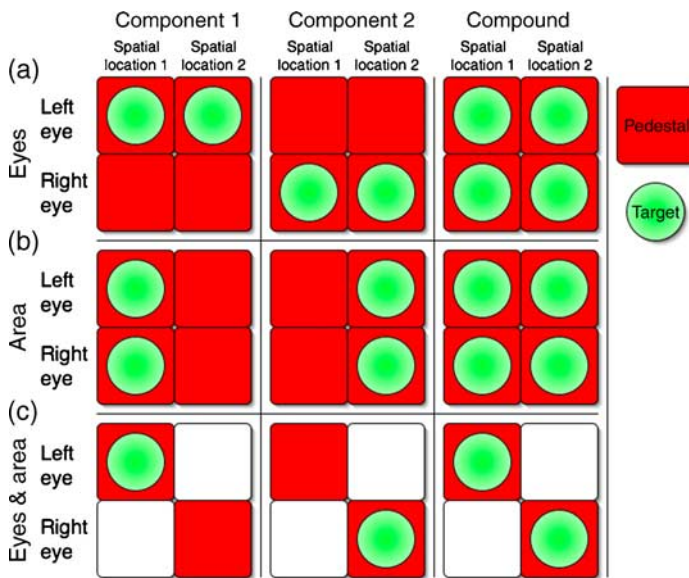


Figure 1. Schematic illustration of the stimulus design logic for investigating contrast summation (a) across eyes (Meese et al., 2006), (b) across area (Meese & Summers, 2007), and (c) across the conjunction of eyes and area (the work here). The filled square (red) and circular (green) symbols represent pedestal and target contrasts, respectively. The different columns denote the component pair and the compound condition for each of the three experiments (a–c). Within each quad of icons, the different rows and columns denote stimuli presented to different eyes and different spatial locations, respectively. In (a), the same target was presented at both spatial locations but to different eyes. In (b), the same target was presented to both eyes but to different spatial locations. In (c), the target was presented to different eyes and different spatial locations. Note that in the experiment here, trials were counterbalanced across location and eye.

the third one as a dual increment. Note that in no condition in Figure 1c is there a pedestal or target increment in the same eye that extends across space, or the same spatial location that goes to both eyes.

### Questions, predictions, and preview of results

Here we ask whether the contrast integration process is generally indifferent to eye of origin, as found by Meese and Summers (2009) at detection threshold. Or will it be that when the experiment is extended above threshold, the emerging sensory evidence for different pedestal stimuli in the two eyes will restrict the target integration to eye of origin? If the first hypothesis is correct, then the dipper function for dual increments should sit above that for single increments across the entire range of pedestal contrasts. If the second hypothesis is correct, then the two dipper functions should converge once the pedestal contrast is above its own detection threshold. We addressed this issue here to further our understanding of the organizing principles of early spatial vision. Our

results are consistent with the first outcome, indicating an area summation process that cares little, if at all, about eye of origin. Our results are well described by a three-stage model of contrast gain control.

## Methods

### Equipment

Stimuli were viewed through Cambridge Research Systems (CRS) ferro-electric (FE-1) shutter goggles and displayed on a Clinton Monoray monitor with a frame rate of 120 Hz using a CRS ViSaGe stimulus generator. The mean luminance of the display viewed through the goggles was 20 cd/m<sup>2</sup>. The shutter goggles allowed different stimuli to be presented to the two eyes by interleaving across frames with a refresh rate of 60 Hz. Using this method, the contrast of the entire stimulus (pedestal plus target) was controlled by look-up tables independently for each eye. Gamma correction was performed to ensure linearity over the full range of stimulus contrasts. Observers sat at a viewing distance of 51.5 cm with their head in a chin and headrest fixating a dark square point (4.8 arcmin) placed in the center of the display throughout the experiment. The experiment was controlled by a PC.

### Stimuli

The two different types of Swiss cheese stimuli are shown in Figure 2. The carrier was a horizontal sine-wave grating in sine phase with the center of the display and had a spatial frequency of 2.5 c/deg. It was contrast modulated (multiplied) by a circular raised cosine function with a central plateau of 8 deg, and a blurred boundary of 1 deg, giving a full-width at half-height of 9 deg. A “raised-plaid” envelope was used to modulate this stimulus further. The plaid was the sum of two sine-wave grating components with orientations of  $\pm 45^\circ$  and a spatial frequency of 0.5 c/deg, each with contrasts of 0.5. This gave minima and maxima of  $-1$  and  $1$ , respectively. The envelope was then “raised” by adding 1 to each point and dividing by 2 throughout, giving minima and maxima of 0 and 1. Thus, the equation for the modulator was

$$env = \left( \frac{1 + \cos(2\pi[fx \cos(\theta) + fy \sin(\theta)] + \phi)/2 + \cos(2\pi[fx \cos(-\theta) + fy \sin(-\theta)] + \phi)/2}{2} \right) / 2, \quad (1)$$

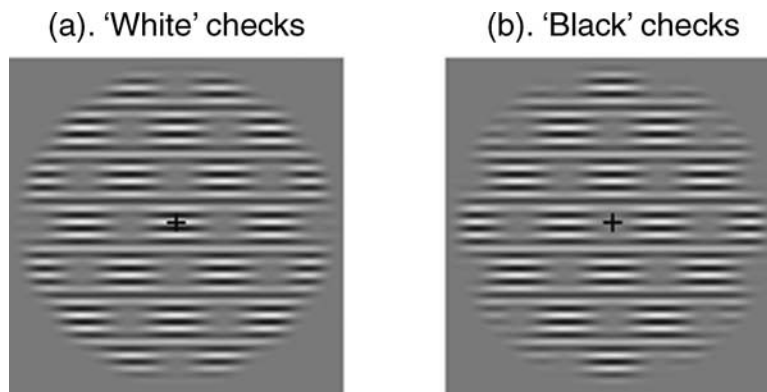


Figure 2. Swiss cheese stimuli (pedestals and targets). One phase of checks (we use the nominal labels of “white” and “black”) was used as the pedestal in one eye, and the other phase of checks was used as the pedestal in the other eye. Contrast increments of the carrier were applied to either the left eye alone, the right eye alone (single increments), or both eyes (dual increment). Stimulus conditions were counterbalanced across eyes. The central crosses here are an aid to free fusion of these sample stimuli (see Discussion section). In the experiments, a small square fixation point was used instead.

where  $f$  is spatial frequency ( $= 0.5$  c/deg),  $\theta$  is orientation ( $= 45^\circ$ ), and  $\phi$  is phase. There were two different phases of modulation: cosine phase ( $\phi = 0^\circ$ ; Figure 2a) and negative cosine phase ( $\phi = 180^\circ$ ; Figure 2b). These stimuli were given the nominal labels of “white” and “black” checks, respectively, as a reference to the magnitude of the modulator at the center of the display (unity and zero).

Note that there are 7.07 cycles of carrier grating for every two checks (i.e., one cycle of a vertical cross-section through the envelope). Note also that the physical sum of the two stimuli in Figure 2 is equal to the carrier grating without the modulation by the raised plaid.

Stimulus contrast is expressed as Michelson contrast in % of the carrier (i.e.,  $c = 100[(L_{\max} - L_{\min})/(L_{\max} + L_{\min})]$ ) or in dB re 1% ( $= 20\log_{10}(c)$ ). Pedestal contrast was controlled by manipulating the carrier contrast ( $c$ ) of the pedestals in each eye (the pedestal contrast was always the same in the two eyes). Target contrast was controlled by applying a contrast increment to the carrier of either one (single increment) or both eyes (dual increment).

Even though the pedestals in the two eyes were very different (see Figure 2), the dichoptic pedestal appeared as a fairly uniform patch of grating with about 24 cycles of the carrier. Although there was sometimes a sense of luster in the stimulus when the target increments were very large (e.g., at the beginning of a session), observers confirmed that this was not the case at and around the contrast increment threshold, where the stimuli appeared uniform.

## Stimulus conditions

The experiment measured contrast-masking functions (dipper functions) over a range of pedestal contrasts from 0% to 32%. The two eyes were always presented with different phases of modulation (i.e., one eye saw “white” checks, while the other eye saw “black” checks; see Figures 2 and 3). In different conditions, contrast increments of the

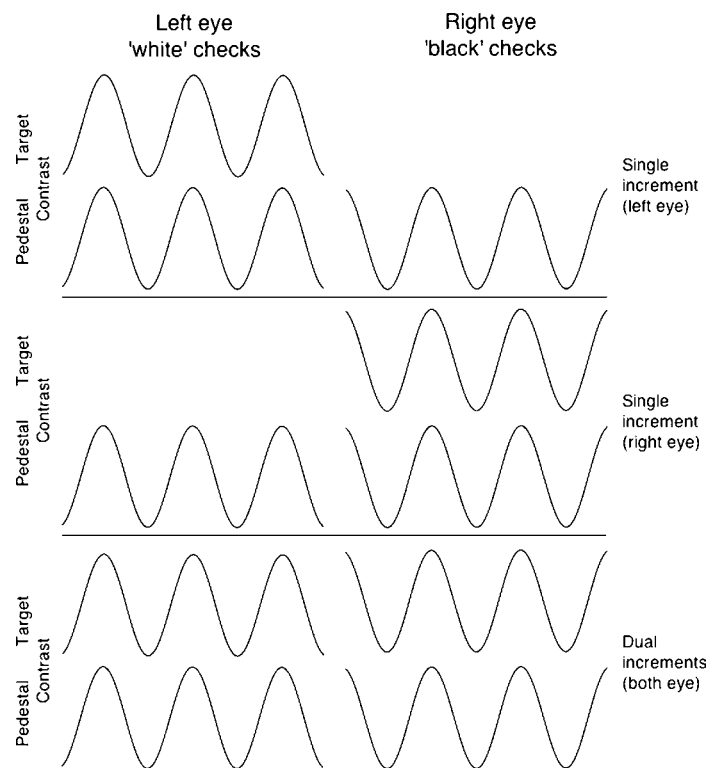


Figure 3. One-dimensional cross-sections of the Swiss cheese contrast modulation (spatial envelopes) for each of the three dichoptic stimulus conditions from Figure 1c (two single increments and one dual increment). The figure shows “white” and “black” checks in left and right eyes, respectively, though in the experiment this was counterbalanced across eyes. Note that the contrast modulation always varied from zero to unity and that the binocular sum of the dichoptic pedestals was a uniform carrier (i.e., there was no destructive interference between the signals in the two eyes). The carrier contrasts (not shown) of the targets and pedestals were varied by the observers (using a staircase method) and experimenters, respectively.

carriers were applied to either (i) the “white” checks only, (ii) the “black” checks only, or (iii) both “black” and “white” checks (Figure 3), and the conditions were counterbalanced across eye. The first two conditions are referred to as “single increments” and the third condition as a “dual increment.” Thus, there were 6 different conditions in total: 2 eyes  $\times$  (2 single increments + 1 dual increment). Our interest was in a comparison between single and dual increments after collapsing across eye.

## Procedure

The level of target contrast (the contrast increment of the carrier) was selected by a three-down one-up staircase procedure (Wetherill & Levitt, 1965) and a single condition was tested using a pair of randomly interleaved staircases (Cornsweet, 1962). The target contrast always began well above detection threshold, and in an initial stage of data collection, a large step size was used (12 dB). After the first reversal, the step size was reduced to 3 dB and data collection continued for a further twelve reversals. These last twelve reversals constituted the test stage for each staircase. We used a two-interval forced-choice (2IFC) procedure where one interval contained only the pedestals and the other contained the pedestals plus target increment(s). The onset of each 100-ms test interval was indicated by an auditory tone and the duration between the two intervals was 400 ms. Observers were required to select the interval containing the target using one of two buttons to indicate their response. Correctness of response was provided by auditory feedback, and the computer selected the order of the intervals randomly. For each run, data from the test stages (above) were collapsed across the two staircases and thresholds (75% correct) and standard errors were estimated by probit analysis (Finney, 1971). Individual estimates for each psychometric function were based on around 100 trials.

Experimental “contrast blocks” were repeated twice. A contrast block consisted of a set of “mini-blocks” for each of ten pedestal contrasts (including 0%). A mini-block consisted of an experimental session for each of the six pedestal and test configurations described above. The order of pedestal contrasts and the order of conditions within each mini-block were determined using a random number generator.

After collapsing across eye, thresholds for each increment type (“black” increment, “white” increment, dual increment) were averaged across four estimates ( $\approx 400$  trials). Staircases started with target contrasts that were sufficiently high for the stimulus condition (“white,” “black,” or dual) to be easily identifiable at the beginning of each mini-block.

Before data collection began (and consistent with much of our earlier work), the following rejection and replacement criterion was set to lessen the impact of unreliable estimates of threshold. If the standard error of a threshold estimate within a mini-block was greater than 3 dB

(estimated by probit analysis), the data for that condition were discarded and the mini-block was rerun.

## Observers

Three undergraduate optometry students performed the experiment as part of their course requirement. They were YR, PP, and ASP. The observers had at least 1 h of practice at the tasks before formal data collection began. All observers wore their normal optical correction and had normal stereopsis, as assessed by random dot stereograms.

## Results

A preliminary analysis of the results confirmed that for all three observers, sensitivity of the two eyes was very similar, and so data were collapsed across eye. Dipper functions are shown for the average of the three observers in Figure 4. Sensitivities to contrast increments of the

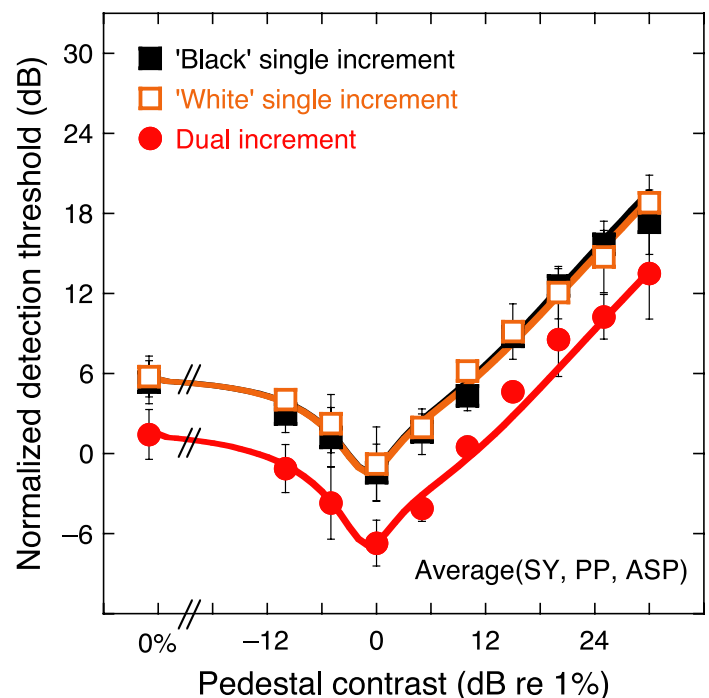


Figure 4. Dipper functions averaged across three observers for each of three conditions ( $\sim 1200$  trials per point). Contrast increments were applied to either “black” checks alone (solid squares), “white” checks alone (open squares), or both “black” and “white” checks simultaneously (solid circles). (See Figure 2 for the meaning of the stimulus terminology.) Error bars show  $\pm 1$  SE across observers. Curves show the fit of the three-stage model of contrast gain control described in the text. The RMS error of the fit was 0.85 dB. See Appendix A for a full list of parameter values. Note that on each axis, a step of 6 dB is equivalent to a factor of 2.



“black” checks (solid squares) and “white” checks (open squares) were very similar (see Figure 2 for the meaning of stimulus terminology). When the contrast increments were applied to both the “black” and “white” checks simultaneously (in different eyes), the form of the dipper function was unchanged but displaced downward (solid circles). The level of contrast summation across the “black” and “white” increments (the contrast differences from Figure 4, expressed in dB) is shown for each of the three observers and their average in Figure 5.

Essentially, contrast summation was found across the conjunction of eyes and space at all pedestal contrasts. There was little or no systematic effect of pedestal contrast, and summation typically fell within the bounds of 3 dB and 6 dB (dashed horizontal lines in Figure 5; i.e., between factors of  $\sqrt{2}$  and 2). Thus, at threshold and above, summation is consistently higher than quadratic (3 dB, or  $\sqrt{2}$ )—a figure often associated with binocular summation (e.g., Campbell & Green, 1965). The high

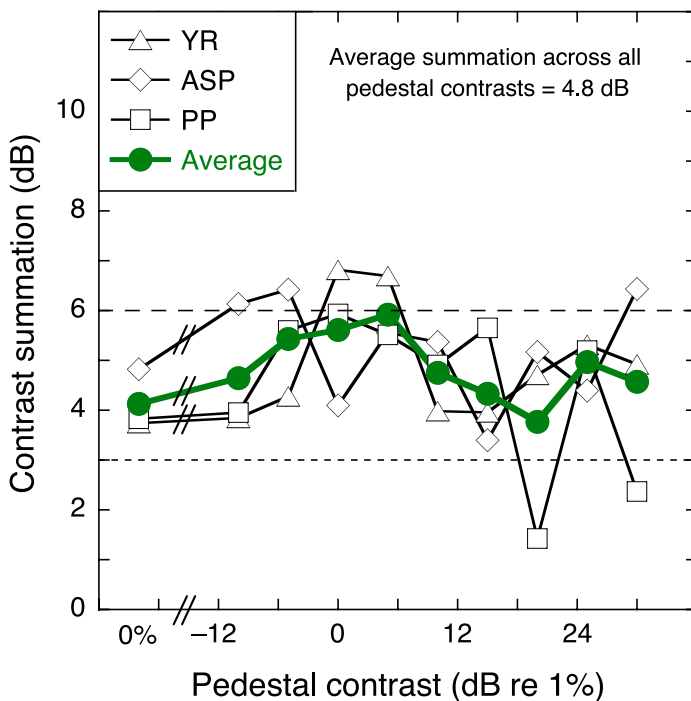


Figure 5. Summation of contrast across eyes and area for each of three observers (open symbols) and their average (solid circles). The fine and coarse dashed lines indicate quadratic summation (3 dB; a factor of  $\sqrt{2}$ ) and perfect linear summation (6 dB; a factor of 2) of the amplitudes of the contrast increments, respectively. Contrast summation for YR and ASP and the average of the three observers were each significantly greater than 3 dB according to a sign test ( $n = 10$ ,  $p < 0.001$ ). This analysis fell slightly shy of significance for PP ( $n = 10$ ,  $p = 0.055$ ). The average summation for a pedestal contrast of 0% was 4.1 dB. The average summation for pedestal contrasts above 0 dB was 4.7 dB. The average summation across the entire dipper function was 4.8 dB. Note that, on each axis, a step of 6 dB is equivalent to a factor of 2.

levels of binocular summation across all pedestal contrasts here (average of 4.8 dB) are only fractionally less than that found in an earlier experiment (5.1 dB), where grating stimuli were spatially superimposed in the two eyes (Meese et al., 2006).

## Three-stage model of contrast gain control

### Model motivation and design

Here we fitted our results with a three-stage model of contrast gain control inspired by previous results on binocular summation and spatial summation at threshold and above. The model builds on the findings and constraints reported in four previous studies (Meese, 2010; Meese et al., 2006; Meese & Summers, 2007, 2009) in a logical and simple way. It extends our previous two-stage model of binocular summation and contrast gain control (Meese et al., 2006) to include the area summation and spatial filtering of Meese (2010) and Meese and Summers (2007). However, in what order should the various stages be placed? Meese (2010) provided evidence at threshold for contrast summation *within* linear filter elements (i.e., individual receptive fields) followed by a nonlinearity and then summation *across* filter elements (i.e., area summation). Meese and Summers (2009) tested 62 formally different models of ocularity and spatial summation against their threshold data and concluded that contrast summation takes place across eyes before area. They also identified transducer-like nonlinearities before and after the binocular summation stage, and also at the output stage. In sum, previous work at threshold suggests an alternating cascade of three transducers and two summation stages all placed after an initial stage of linear filtering (Meese, 2010; Meese & Summers, 2007, 2009). Here, we retain the two stages of summation and the three stages of contrast transduction but raise the application of the model above threshold by converting the transduction stages to contrast gain control. This is consistent with the gain control in the pre- and post-binocular summation stages of Meese et al. (2006) and the gain control at the spatial summation stage of Meese and Summers (2007). It is also broadly consistent with ideas about contrast normalization (divisive inhibition) at the single-cell level (e.g., Heeger, 1992; Kouh & Poggio, 2008; Ringach, 2010). We will consider the purpose of the first two stages of gain control in the next subsection and in part 3 of the discussion. The third stage of gain control (the output stage) is needed to fit quantitative aspects of the dipper functions (see next section but one). However, as we have argued before (Baker, Meese, & Georgeson, 2007; Meese et al., 2006), this final stage is probably best treated as a numerical convenience rather than an explicit process of divisive inhibition, at least until it is better understood (see also the [Internal noise in the model](#) section and [Appendix B](#)).

The decisions about where to put the nonlinear transducers, the gain controls, and the summation stages were all driven by previous results, as described above. However, there are some model details for which our existing data do not provide strong constraints (e.g., pooling over filter phase at Stage 2). Thus, the specific model shown here is done in the spirit of demonstrating a plausible visual arrangement that can account for our results. Nonetheless, after trying many different variants of the model, we were unable to find a simpler one that maintained all of the properties of the current one. We provide an overview of two of these properties next.

### Overview of two key model properties

The first stage of gain control is placed before/at the binocular summation stage and is needed to normalize sensitivities of the two excitatory monocular channels to achieve *ocularity invariance*—the property that a pattern’s suprathreshold contrast looks similar with one eye and two (Baker, Meese, & Georgeson, 2007). The second is placed before/at the area summation stage and normalizes sensitivity to different areas to achieve *area invariance*—the property that a pattern’s suprathreshold contrast looks similar when it changes in size (Cannon & Fullenkamp, 1991a; Meese & Summers, 2007). The third stage is needed to refine the fits to the dipper functions but does not affect summation across eyes or space. We will elaborate on these requirements in the discussion.

An overview of the model parameters and other model details follows, but readers could skip to the discussion without loss of continuity. A schematic outline of the model is shown in Figure 6, and full details are described in Appendix A.

### Model parameters: Three fixed and five free

The spatial filters were those used by Meese and Summers (2007). They were Cartesian separable log Gabor filters matched to the spatial frequency and orientation of the carrier (see Meese, 2010 for the equation). The initial filter response exponent ( $m$ ) was set to 1.2, consistent with previous studies (Meese et al., 2006; Meese & Summers, 2009) and so that the level of conventional binocular summation is greater than quadratic ( $m = 2$ ) but less than perfectly linear ( $m = 1$ ). This is followed by the Stage 1 gain control of Meese et al. (2006;  $S = 1$ ) and then binocular summation. The analysis of Meese and Summers (2009) demanded an intermediate exponent ( $r$ ) between these stages. This was set to 1.67 here so that the overall exponent before area summation was  $mr = 2$ , consistent with energy detection and previous models of psychophysical results (Klein & Levi, 2009; Meese, 2010; Meese & Summers, 2009; Petrov, Doshier, & Lu, 2005) and single-cell physiology in V4 (Ghose & Maunsell, 2008). This second stage of gain control is accompanied by area summation. For convenience, this

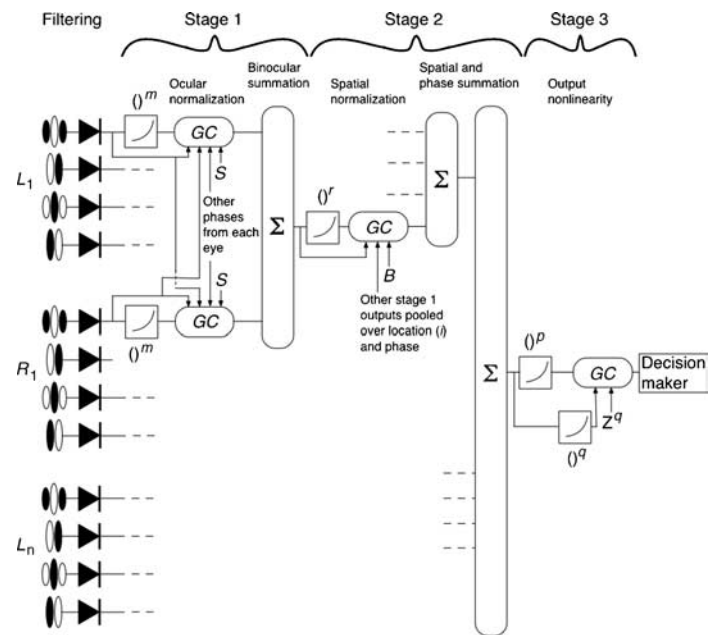


Figure 6. Schematic outline of the three-stage contrast gain control model used for the experiment here (the fit is shown in Figure 4). Arrows denote divisive inhibition. Not shown is (i) the attenuation surface that precedes the spatial filtering and represents retinal inhomogeneity (see Appendix A for details) and (ii) the location of performance limiting noise, which is discussed in the text. Abbreviations and parameters: GC, gain control,  $L_i$ ,  $R_i$ , left and right eye contrasts at the  $i$ th corresponding retinal point;  $S$ ,  $B$ ,  $Z$ , saturation constants;  $m$ ,  $r$ ,  $p$ , excitatory transducers;  $q$ , suppressive transducer;  $\Sigma$  linear summation of input lines. See Appendix A for full list of parameter values.

was performed over the entire stimulus region, though the true spatial extent might be somewhat less than this (Meese, 2010; Meese & Summers, 2007, 2009).

The exponents ( $p$  and  $q$ ) at the third stage of gain control are used as a final control of the steepness of the log–log slope of the upper limb of the masking function (the dipper handle). The exponent  $p$  also influences the shape and position of the “dip.”

Stage 3 is followed by performance limiting additive Gaussian noise with standard deviation  $\sigma$ . If the model were stochastic, then the observer (the “decision maker” in Figure 6) would choose the 2IFC interval that produced the greater response. In the deterministic version of the model (used here), threshold is the contrast needed for the response difference across 2IFC intervals to equal  $k$ . This parameter is linearly related to  $\sigma$  and does not represent an additional degree of freedom in the model. (We shall return to the placement of noise in the section after next.)

Overall, there were five degrees of freedom in the model. In general, four parameters control the basic shape and position of dipper functions (Legge & Foley, 1980). The equivalent parameters here were  $Z$ ,  $k$ ,  $p$ ,  $q$ . The fifth



free parameter ( $B$ ) was needed to condition the second stage of gain control. These parameters have little or no effect on the level of summation, which is an emergent property of the model's architecture. The fixed parameters were  $m$ ,  $r$ , and  $S$ , plus the filter parameters, which were not critical in this study.

In a refinement to the model, we arranged that the saturation constant of the second stage of gain control ( $B$ ) was suppressed by the excitatory component of that stage (and this is the version of the model used in Figure 4). This had the advantage of improving the area invariance property of the model without introducing extra parameters (see forward to Figure 7 in the discussion). However, we were able to achieve a good fit to the data here without this refinement in place (RMS error = 1.0 dB; not shown).

The five parameters were adjusted by a simplex algorithm to fit the averaged data (Figure 4). The RMS error of the fit was 0.85 dB. Note that model curves and data are very similar for the two different phases of check ("black" and "white"), and that these are displaced upward from the dual increment condition.

### Phase interactions

For completeness, we have included four phases of filter (with 90° offsets between them). We pool across filter phase in the suppression pathway at Stage 1, consistent with the analysis of dichoptic masking performed by Baker and Meese (2007). We have included excitatory summation across filter phase at Stage 2 but before spatial pooling, as this allows the higher order mechanisms to retain the bandwidths of the early mechanisms (Wilson, Wilkinson, & Asaad, 1997). However, the locus of phase pooling across half-wave rectified filters (Baker & Meese, 2007; Georgeson & Meese, 2007; Read, Parker, & Cumming, 2002) is not well constrained by the experiments here and the details of this stage might need to be revised in the light of further evidence.

### Internal noise in the model

In the implementation of the model here, we placed the limiting additive noise just before the decision maker. This late position was motivated by Birdsall's theorem (Klein & Levi, 2009; Lasley & Cohn, 1981), which states that performance-limiting noise will "linearize" subsequent monotonic nonlinearities (such as the contrast transduction at Stage 3) when measuring performance. This poses a problem because other work concludes that the limiting noise for *contrast detection* is placed after the exponent product  $mr$  but *before* area summation (Meese, 2010; Meese & Summers, 2007, 2009). However, the final stage of gain control can be interpreted in several ways (Baker, Meese, & Georgeson, 2007; Meese & Summers, 2009). For example, it might not

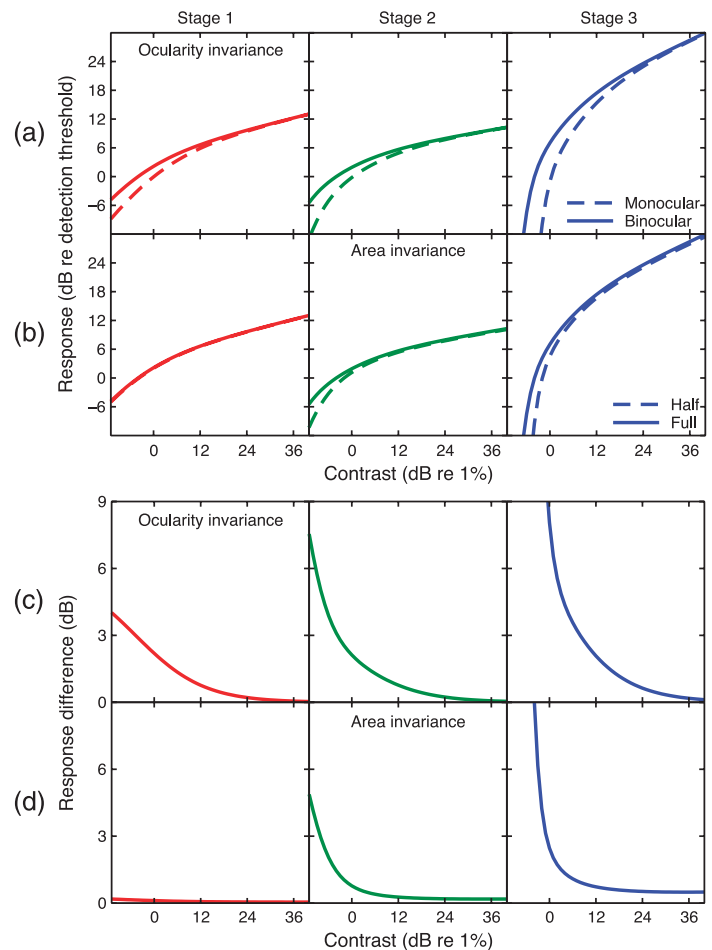


Figure 7. Behavior at the output of each of the three stages in the model. (a) Contrast responses to monocular and binocular gratings. The responses are arbitrarily normalized to the response for a full binocular grating at contrast detection threshold. (b) Responses to two different sized gratings (the "half" grating was a full grating halved down the middle). (c) Difference (in dB) between the two curves in (a). (d) Difference (in dB) between the two curves in (b). Ocularity invariance is achieved for moderate contrasts and above at the output of Stage 1, but at low contrasts, there is a binocular advantage. Area invariance is achieved for moderate contrasts and above at the output of Stage 2, but at low contrasts, there is an area advantage. There is no area advantage at the output of Stage 1 as this precedes area summation.

represent transduction and divisive suppression, but intrinsic uncertainty at threshold and the loss of uncertainty associated with the pedestal above threshold (Meese & Summers, 2009; Pelli, 1985). Multiplicative noise might also be relevant at this stage, particularly above threshold (Georgeson & Meese, 2006; Kontsevich, Chen, & Tyler, 2002; Lu & Dosher, 2008). However, for the sake of brevity and simplicity, we have not attempted to develop such model variants here—the arrangement in Figure 6 with late noise is perfectly adequate for the results in Figure 4 (see also Appendix B).

### Model behavior for other pedestal experiments

As described earlier, our aim was to produce a model that could fit the data here but was motivated by previous results. Therefore, we might hope that our three-stage model of contrast gain control will also provide a good account of other relevant data sets. In [Appendix B](#), we show that it is consistent with the results from Legge and Foley (1980), Meese (2004), Meese et al. (2006), and Meese and Summers (2007).

### Stereopsis

We note that architecture related to our proposal ([Figure 6](#)) might be extended for involvement in stereopsis (Read et al., 2002). For example, summation across space and phase is withheld until Stage 2, so appropriately wired outputs at Stage 1 could provide an appropriate feed for local disparity detectors. However, binocular convergence is not limited to the processes involved in stereo depth (e.g., Peirce, Solomon, Forte, & Lennie, 2008) and here we shall restrict our inferences and discussion to the nonstereo domain of spatial vision.

## Discussion

We shall summarize what we have found and then précis each of the three main parts of the discussion in reverse order.

We have demonstrated that the visual system is able to combine image contrasts from different positions in the two eyes across a wide range of base (pedestal) contrasts. However, this counterintuitive result does not require mechanisms that are specialized for this situation to explain it. In essence, the model that we propose ([Figure 6](#)) performs conventional binocular summation followed by summation of a binocular representation that is fairly uniform over space. This area summation stage is simply indifferent to which eye the different parts of the surface originated (Hess & Field, 1995; Huang, Hess, & Dakin, 2006; Mansouri, Hess, Allen, & Dakin, 2005; Meese & Summers, 2009). Thus, we see our laboratory-based result as an emergent property of the simple visual architecture that we propose and not one that reveals a particular computational goal or ecological need regarding dichoptic stimulation. In the third part of the [Discussion](#) section, we consider the general issue of why the system might be arranged to achieve the extensive spatial integration that we propose.

The second part of the [Discussion](#) section presents a critical appraisal (of design issues), outlining why we but not others (e.g., Chirimuuta & Tolhurst, 2005; Legge & Foley, 1980) have found suprathreshold summation of luminance contrast. However, a central component of the study here involved presenting different stimuli to different

eyes and so we first briefly consider some technical details of the long-standing issues of binocular fusion and rivalry.

### Part 1: Fusion and binocular rivalry

Our conclusion regarding summation across area requires that the dichoptic stimuli were presented to corresponding retinal regions. If, for example, the two images were horizontally misregistered by as much as one check width ( $\sqrt{2}$  deg, equivalent to 3.5 carrier cycles), then a substantial part of the stimulus would be binocular checks, and increment summation would not be surprising (Meese et al., 2006). However, in the experiments, we used stereo goggles, which meant that the point of convergence and accommodation were the same. This provides natural viewing conditions and permits binocular fusion with ease, particularly since we used a clearly visible dark fixation point in the center of each eye's image. Furthermore, even though we presented different stimuli to the two eyes, the cyclopean image appeared as a uniform grating, at least when the pedestals were detectable and the target contrast was close to increment threshold (i.e., the contrasts in the two eyes were similar; see [Methods](#) section). This can be verified by free-fusing the “black” and “white” checks in [Figure 2](#). So long as steady fixation is maintained, the stimulus will appear as a plain unmodulated grating. (This might take some practice but is compelling when it is achieved. We found best results using a short viewing distance [30–50 cm], converging in front of the picture plane and holding fixation very steady.) Observers reported that this is how the stimulus appeared during the experiment. If the images had been misregistered, then low-contrast regions would have been visible. Thus, we think that binocular misregistration of our dichoptic stimuli is not a critical factor in these experiments.

A striking phenomenon that can arise when different stimuli are presented to different eyes is binocular rivalry. However, we also doubt that this was important here. First, we used brief presentation times (100 ms), which are probably too short for rivalry to initiate. Second, although the images in the two eyes were different, the carriers were identical. Thus, there was no conflict of local features, other than their contrast, and evidence elsewhere shows that mismatched dichoptic contrast is combined effectively within the scheme that we propose (Baker, Meese, & Georgeson, 2007). Third, the images appeared as continuous gratings (as described above), suggesting that rivalry was not a serious factor. Finally, our model includes a pathway for interocular suppression (see [Appendix A](#)), which contributes to ocularity invariance (Baker, Meese, & Georgeson, 2007; Meese et al., 2006) and dichoptic masking (Baker & Meese, 2007; Legge, 1979), is easily extended to include interocular cross-oriented influences (Baker & Meese, 2007; Baker, Meese, & Summers, 2007; Meese & Baker, 2009), and presum-

ably operates in primary visual cortex (Moradi & Heeger, 2009). With additional temporal dynamics (not considered here), this might also be responsible for at least some part of binocular rivalry (Baker & Graf, 2009; Pearson, Tadin, & Blake, 2007; Van Boxtel, Van Ee, & Erkelens, 2007; Wilson, 2003). In other words, our model includes architecture with sufficient flexibility to make it consistent with current models of eye-based rivalry (see Baker & Graf, 2009; Baker, Meese, & Summers, 2007 for further discussion) and does not pose an obvious challenge to that work.

## Part 2: Design issues

The result here (suprathreshold summation of contrast across eyes and space) is perhaps surprising in the light of many previous studies where either empirical evidence for suprathreshold summation has been lacking (Legge & Foley, 1980; Legge, 1984; Maehara & Goryo, 2005; Meese, 2004), the levels of summation have been small (Bonneh & Sagi, 1999; Meese, 2004), or variability across observers has made interpretation difficult (Meese, 2004; Meese et al., 2005). However, as we describe below, there were several aspects of our stimulus design that we think were important for revealing a robust summation process.

### Factor 1: Retinal inhomogeneity

It is well known that the retina is inhomogeneous, becoming less sensitive away from the fovea (Foley et al., 2007; García-Pérez, 1988; Pointer & Hess, 1989; Robson & Graham, 1981; Wilson & Bergen, 1979). For experiments that extend the size of the stimulus by increasing the area of a centrally placed patch of grating (e.g., Foley et al., 2007; Meese & Summers, 2007; Rovamo et al., 1993), this compromises the level of summation expected and the power of the experiment. This problem with retinal inhomogeneity does not arise with the stimuli used here: a compound (the linear sum) of the two components in Figure 2 is a stimulus with greater “contrast area”<sup>2</sup> but has the same diameter as each of its components. The modeling by Meese and Summers (2007) considered sensitivities to the two stimuli in Figure 2 and confirmed that they should be equally detectable by a process that sums over area, regardless of whether retinal inhomogeneity was included in the model (see also Meese, 2010). The empirical finding that sensitivity is the same for the “black” and “white” versions of the Swiss cheese stimuli in Figure 2 confirms this expectation.

### Factor 2: Noise

System components, be they biological or otherwise, are inherently noisy. This means we should expect early

noise (that which arises before neuronal convergence), to propagate through to the decision variable (Campbell & Green, 1965; Kontsevich & Tyler, 1994; Tyler & Chen, 2000; Watt & Morgan, 1984). This is important for conventional area summation experiments where the diameter or width of the target grating is varied (Foley et al., 2007; Meese et al., 2005; Robson & Graham, 1981; Rovamo et al., 1993) because the noise level will also increase with area (see Figure 6). Whether this will affect performance depends on whether the dominant source of noise is early or late. Conclusions on this have differed with respect to stimulus area (Foley et al., 2007; Meese, 2010; Summers & Meese, 2007) and other factors (Bowne, 1990; Itti, Koch, & Braun, 2000), though our own work suggests that it is placed before area summation (Meese, 2010; Meese & Summers, 2007). However, we doubt that this presents a problem for the experiment here. Our stimulus design prompts the simplifying assumption that the visual system processes inputs from the same retinal regions for all three stimulus conditions (the two single increment stimuli, and the dual increment stimulus; see also Meese, 2010). From this, it follows that the neuronal noise sources are also the same across the conditions, regardless of their placement in each model. The alternative is to suppose that the observer can construct a template similar to the target region (the Swiss cheese modulator), and down-weight the contribution of noise (and signal) from the less informative low-contrast regions of the target. One way to investigate whether the visual system employs this strategy would be to derive the classification images (e.g., Solomon, 2002) for the stimuli in Figure 2. Another method is to perform an identification task at threshold (Watson & Robson, 1981). Meese and Summers (2007) did this for binocular stimulation. They found that observers could not distinguish between a Swiss cheese stimulus (Figure 2a) and a full stimulus (i.e., a carrier without “holes”; the sum of Figures 2a and 2b) for full field pedestal contrasts of either 0% or 25% when the contrast increments were equally detectable. This makes it seem unlikely that the observer uses a template of the Swiss cheese when performing this task, and therefore, we doubt that our conditions were confounded with (substantially) different levels of noise.

### Factor 3: Region of suppression

A third aspect of the design was that for each pedestal contrast, the pedestal structure for the three stimulus conditions was the same (opposite phase Swiss cheese in each eye). This was intended to hold the overall level of suppression approximately constant across conditions, including interocular suppression (Baker & Meese, 2007; Baker, Meese, & Georgeson, 2007; Ding & Sperling, 2006; Kontsevich & Tyler, 1994; Meese et al., 2006) and lateral suppression (Cannon & Fullenkamp, 1991b; Meese, 2004; Tolhurst, 2007; Xing & Heeger, 2000). As



recognized elsewhere (Bonneh & Sagi, 1999; Meese, 2004; Meese et al., 2005; Meese & Summers, 2007), it is important that this is done so that the target region is not confounded with (substantially) different levels of suppression.

#### Factor 4: Region of excitation

As we have just described, our experiment was designed to try and clamp the contrast gain control, but this is not necessarily sufficient to reveal the summation process. If the region of suppression is fixed, but the observer is able to adjust the region of excitatory integration to (approximately) match the target (Meese, 2004), then, even when the noise is fixed (see above), model behavior can be counterintuitive: its performance can even decline as the area of the signal grows (see Meese, 2004). This potential problem is avoided if the observer's integration regions of excitation and suppression are *both* held constant. As discussed above in the context of noise, we believe this is achieved by the stimuli here: that excitatory integration takes place over a constant retinal diameter regardless of whether the target is a dual or single increment. This is important for two reasons. First, it is likely to invoke the use of the same detecting mechanism (or mechanisms) for all the stimuli here, and therefore the same region of suppression (see [Factor 3: Region of suppression](#) section). Second, with this arrangement, it is arguable that the summation process is revealed most effectively because the benefit of signal integration is accompanied by a release from *dilution masking*. Dilution masking is a theoretical process that occurs when uninformative stimulus regions (i.e., pedestal without target) are summed with informative stimulus regions (i.e., pedestal plus target) on the numerator of the contrast gain control equation. It is a different process from each of those involved in conventional within-channel (Legge & Foley, 1980) and cross-channel (Foley, 1994) forms of masking. When the target covers the entire pedestal, dilution masking cannot operate—there is no pedestal region *without* target—and performance improves, partly for this reason (see Meese & Summers, 2007).

#### Design conclusions: A novel approach

The average level of model summation across all pedestal contrasts ([Figure 4](#)) was 5.32 dB (a bit less than a factor of 2). This is slightly greater than the 4.8 dB found in the experiment (though note the large variation across contrasts and observers in [Figure 5](#)), suggesting that the true summation process is fractionally less effective than in our model. This could be because (1) the spatial extent of summation was somewhat less than two full checks of the modulator for the conditions here, (2) that observers were able to perform some restricted exclusion of some of the factors described above in the single increment conditions (e.g., by using second-order mechanisms<sup>3</sup>), or (3) that very small amounts of vertical image misregistration of the

carriers occurred. Nevertheless, we think that the experiment here and those in Meese et al. (2006) and Meese and Summers (2007) were successful in revealing substantial levels of suprathreshold summation because (to a first approximation) we avoided confounding influences from each of the following factors: (i) retinal inhomogeneity, (ii) internal noise, (iii) contrast gain control from the pedestal, and (iv) the region of summation. We know of no other attempts to prevent all four of these parameters from co-varying with the independent variable of target area.

### Part 3: What are the early stages of contrast vision doing?

#### Contrast integration across eyes and space: Signal combination or signal selection?

We have performed several previous experiments and analyses at detection threshold that are closely related to the suprathreshold experiments here (Meese, 2010; Meese et al., 2006; Meese & Summers, 2007, 2009). In all cases, we concluded that summation effects were too large to be attributed to probability summation or a MAX operator (signal selection) but were consistent with a signal combination process. Unfortunately, similar formal analyses are less powerful above threshold because (i) the slope of the psychometric function is typically quite shallow for contrast discrimination (e.g., Weibull  $\beta \sim 1.3$ ; Meese et al., 2006) and this means the data are less well constrained by competing models (see Meese & Summers, 2007, 2009), and (ii) results tend to be more variable above threshold (e.g., see [Figure 5](#)) owing, at least in part, to the shallow psychometric function (not shown). Nevertheless, in our studies the typical levels of summation above threshold are similar (or a little greater) than those at threshold, and gain control models involving signal combination for the entire contrast range provide a good, parsimonious account of the results (Meese et al., 2006; Meese & Summers, 2007; [Figure 4](#)). Given the need for neuronal convergence and signal combination in models of pattern and object recognition (e.g., Cadieu et al., 2007) and the evidence for it from single-cell physiology (Pollen, Przybyszewski, Rubin, & Foote, 2002), this conclusion is perhaps not surprising. Nonetheless, it stands in the wake of a substantial body of visual psychophysics that has been influenced by the first and second dogmas of spatial vision (see [Introduction](#) section) and has supposed otherwise.

We have not attempted a formal assessment of the spatial extent of signal combination here, though the relation between the spatial frequencies of the carriers and modulators that we have used suggests that it extends over at least 7 carrier cycles, equivalent to two checks of the modulator (see Meese & Summers, 2007). Other work at detection threshold suggests a minimum of 16 cycles (Meese, 2010). No doubt, this figure will continue to be revised in the light of further work.

### A possible heuristic for spatial integration

Meese and Summers (2009) performed similar experiments to those performed here but restricted the investigation to threshold (i.e., pedestal contrast = 0%). In that study, we concluded that binocular summation precedes spatial summation in the system hierarchy (Figure 6; see also Mansouri et al., 2005). Luminance contrasts in common spatial frequency and orientation bands are summed across eyes from corresponding retinal points to produce a binocular image. For the results there and here, our model proposes that this is followed by widespread integration of contrast across area. Further work is needed, but a system that does this with complete disregard for underlying image statistics or structures seems unlikely. A more plausible visual heuristic is to integrate over spatial regions for which the local analysis is similar, or for which the carrier and/or modulator change only smoothly in the binocular image. An adaptive or matched filtering process such as this might involve local comparisons of image structure (Field, Hayes, & Hess, 1993; Kingdom, Prins, & Hayes, 2003; Levi, Klein, & Chen, 2005; Saarinen, Levi, & Shen, 1997) and spectrum (Abbey & Eckstein, 2007; Georgeson & Meese, 1997, 1999; Meese & Georgeson, 2005) under the control of second-order binocular mechanisms (Dakin & Mareschal, 2000; Graham & Sutter, 1998) that assess the envelope (Georgeson & Schofield, 2002) and the carrier (Kingdom et al., 2003; Motoyoshi & Kingdom, 2007; Motoyoshi & Nishida, 2004) to search for boundary cues (Grigorescu, Petkov, & Westenberg, 2004; Kawabe & Miura, 2004; Sillito, Grieve, Jones, Cudeiro, & Davis, 1995) or other Gestalt-like cues (Sayim, Westheimer, & Herzog, 2010), plausibly in V2 (Anzai, Peng, & Van Essen, 2007; Mareschal & Baker, 1998), with integration taking place at a later stage, plausibly V4 (Arcizet, Jouffrais, & Girard, 2008; Desimone & Schein, 1987; Pollen et al., 2002) or IT/LO (Köteles, De Mazière, Van Hulle, Orban, & Vogels, 2008; Ostwald, Lam, Li, & Kourtzi, 2008). For stimuli such as those found here, where the carrier is constant and the contrast modulation is not detected at the detection threshold of the target (Meese & Summers, 2007), this heuristic would demand blanket integration over the carrier, consistent with our results. This is equivalent to constructing a (phase-insensitive) template that is matched to the sum of the two stimulus components in Figure 2 by summing lower order (V1-like) filter elements (Rovamo et al., 1993; Watson & Ahumada, 2005). This type of neuronal convergence could be a plausible first step to solving the binding problem for spatially extensive textures (Arcizet et al., 2008; Cant, Arnott, & Goodale, 2009; Graham & Sutter, 1998; Köteles et al., 2008; Roach, Webb, & McGraw, 2008; Webb, Roach, & Peirce, 2008), depth gradients (Meese & Holmes, 2004; Summers & Meese, 2006), and other smooth image structures (May & Hess, 2007a, 2007b). Of course, under normal viewing conditions, images of the natural world tend to have more similarities than differences between the

two eyes, and so perhaps it is not surprising that eye of origin is largely irrelevant within the scheme that we have proposed (Figure 6).

### Ocularity invariance, area invariance, and texture/pattern coding

The perceived contrast of a herd of zebra does not change with the arrival of more zebra (Cannon & Fullenkamp, 1991a, 1991b; Legge & Foley, 1980; Meese et al., 2005) or when you close one eye (Baker, Meese, & Georgeson, 2007; Ding & Sperling, 2006; Legge & Rubin, 1981; Meese et al., 2006; Moradi & Heeger, 2009). We use the terms *area invariance* and *ocularity invariance* to refer each of these, respectively. However, these invariances are at odds with what is needed at detection threshold, where the system would benefit from signal summation. Our specific proposal (Figure 6) has the benefit of meeting each of these demands. Binocular and spatial integrations enhance performance for weakly visible stimuli, but contrast normalization by suppression (gain control) achieves ocularity invariance (Stage 1) and area invariance (Stage 2) for contrasts well above threshold (see Figure 7). This might help explain why it has been easier to demonstrate suprathreshold spatial pooling for tasks in which the dependent variable was not contrast (Dakin & Bex, 2002; Dickinson & Badcock, 2007; Jones, Anderson, & Murphy, 2003; Levi & Klein, 2000; Mareschal, Morgan, & Solomon, 2010; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Wilkinson, Wilson, & Habak, 1998; Wilson & Wilkinson, 1998; Wilson et al., 1997) than when it was (Kersten, 1984; Legge & Foley, 1980; Levi & Klein, 2000; Morrone, Burr, & Vaina, 1995; Nachmias, 2002).

Finally, note how the contrast response (signal transduction) becomes steeper in Figure 7a as the signal progresses through the model from left to right owing to the cascading gain controls. A similar effect has been observed at the single-cell level, where response acceleration becomes more pronounced moving from the LGN through V1 to MT (Sclar, Maunsell, & Lennie, 1990; see also Priebe & Ferster, 2008).

### Why sum, when the benefits are cancelled by countersuppression?

The unusual stimulus configurations used here and elsewhere (Meese et al., 2006; Meese & Summers, 2007) were designed to reveal the suprathreshold operation of contrast summation across eyes and space. However, for more conventional suprathreshold stimulus arrangements—such as when the diameter of a grating is varied—there is no sensitivity benefit with grating area because of countersuppression from the gain controls (Legge & Foley, 1980; Legge, 1984; Meese et al., 2006, 2005; see Part 2: Design issues section). This raises the

question of what the summation process is intended to achieve under normal viewing conditions. That is, why bother to sum suprathreshold stimuli in the first place if this operation is to be effectively nullified by counter-suppression? One possibility (rather different from the adaptive approach described in the previous section) is that populations of variously sized contrast-energy mechanisms (Meese, 2010; Meese & Summers, 2009) encode the spatial extents of luminance contrast patterns (Figure 8). This could be achieved without confounding the pattern's contrast code (Heeger, 1992) if there were a sufficiently large suppression field common to each mechanism in the population. Figure 8 schematizes the population responses to four grating diameters at a contrast well above threshold. The four mechanisms (green ellipses) pool excitatory inputs from many classical receptive fields (small blue ellipses) but from different sized regions of the retina. The four pooling mechanisms respond equally well to the tiny stimulus (see asterisks in Figure 8b), but as stimulus diameter increases, the inhibition from the large suppression field (red) wins out against the mechanisms whose excitatory pooling fields

are too small to benefit from extra excitation. Thus, the pattern of responses across the population varies with stimulus size (as required), but the peak strength of response increases only with stimulus contrast (Cannon & Fullenkamp, 1991a; Meese et al., 2005).

### Model predictions for the physiology

Our scheme makes two intriguing sets of predictions. First, the net contrast response of the population should decrease with an increase in stimulus size once it exceeds that which excites the smallest receptive field (Figure 8). In fact, there is support for this prediction from fMRI (Nurminen, Kilpeläinen, Laurinen, & Vanni, 2009; Press, Brewer, Dougherty, Wade, & Wandell, 2001; Williams, Singh, & Smith, 2003).

The second prediction has several parts and concerns single-cell activity. When measuring receptive field size using a single small oriented bar (or small patch of grating), we predict that (1) various sized receptive fields should be found (including large ones) at both threshold and above. However, when varying the number of bars in

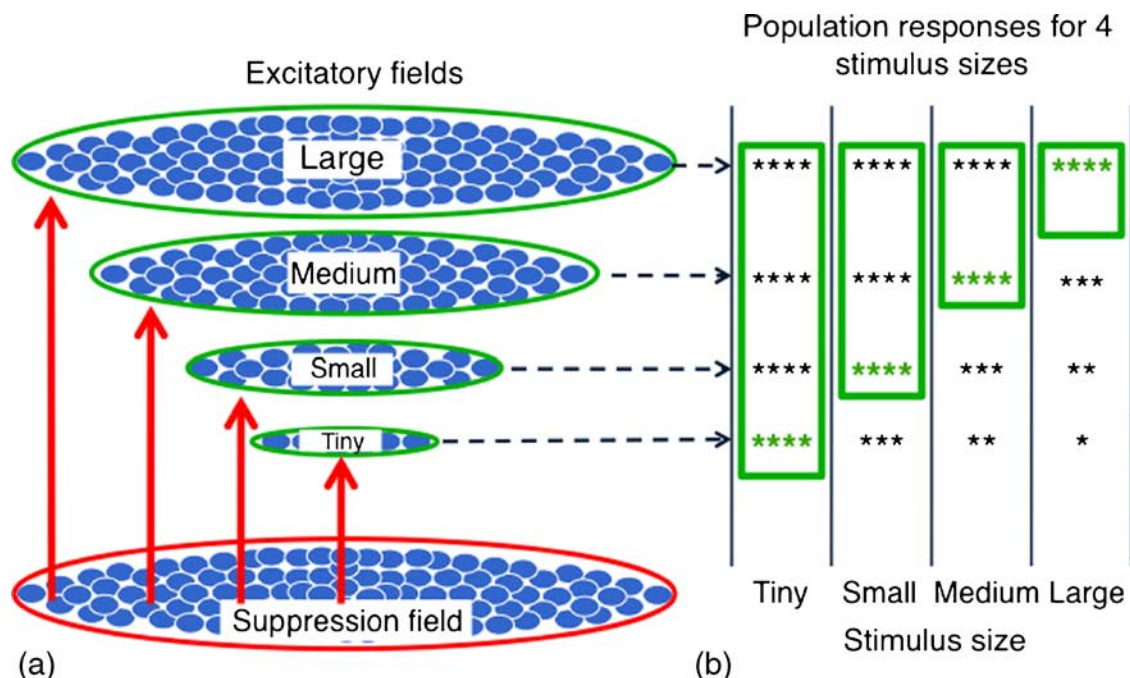


Figure 8. A population code for suprathreshold contrast area. (a) Different cortical contrast mechanisms pool over many classical receptive fields from various sized regions on the retina (green ellipses) but are suppressed by contrast from a common large region (red ellipse). (b) The response of each mechanism decreases with stimulus size (left to right) once the stimulus exceeds the size of its excitatory region (asterisks denote response levels). The response of each mechanism increases with stimulus contrast (not shown), but the distribution of activity across the population (asterisks) does not. This provides the basis for a contrast invariant size code. For example, stimulus size is signaled by the label of the smallest excitatory mechanism that has the strongest response in the population (green asterisks). The suppression field becomes effective only at moderate contrasts and above which means that stimulus size does improve performance around detection threshold (not shown). However, for a contrast discrimination task above threshold, there is little or no performance benefit from increasing stimulus diameter because of the counteracting effects of suppression (Meese et al., 2005). Note that this restriction was not found for the experiments here because the Swiss cheese stimuli (Figure 2) allowed us to change contrast area without changing the stimulus diameter (see Part 2: Design issues section).



a grating, (2) various sized receptive fields should be seen only at low contrasts (around detection threshold). At higher contrasts, as the number of bars is increased we should expect (3) strong lateral inhibition (in the spatial domain) for the cells with small excitatory regions, and (4) no change in response for cells with large excitatory regions. As far as we know, nobody has set out to test this hypothesis and make the appropriate measurements from the same population of cells. However, Desimone and Schein (1987) and Pollen et al. (2002) did find cells in V4 that appear to conform to properties 1, 3, and 4. In addition, several studies have reported that receptive field sizes of some striate cells appear to shrink at higher contrasts, which might relate to the third property (e.g., Cavanaugh, Bair, & Movshon, 2002; Kasamatsu, Miller, Zhu, Chang, & Ishida, 2010).

### Further comments

More generally, if the texture mechanisms we propose (Figure 8) were also tuned to various spatial patterns (e.g., radiating or concentric contours; Ostwald et al., 2008; Wilson & Wilkinson, 1998) and the suppression field pooled over a wide range of pattern elements (e.g., different orientations), this scheme could form the basis of a general-purpose module for encoding visual texture, pattern, and form (Gallant, Conner, Rakshit, Lewis, & Van Essen, 1996).

In spite of our conclusions about signal combination, we do not claim that signal selection (e.g., the MAX operator, or its approximation) is not involved in the visual hierarchy. We expect, however, that it will be revealed by psychophysical experiments that tap into signal invariances (Cadieu et al., 2007; Riesenhuber & Poggio, 1999) that are not also involved in signal combination. The invariant responses of object and face mechanisms with viewing angle might be a good example: there is no obvious benefit to performing signal combination over multiple viewpoints. Remarkably though, the equations and neuronal circuitry needed for the conceptually very different operations of signal combination and signal selection might be quite similar, differing only in the detailed choice of parameter values (see Kouh & Poggio, 2008).

## Summary and conclusions

For 30 years, models of spatial contrast vision have been constrained by the view that spatial pooling of luminance contrast is absent for contrasts above threshold (Legge & Foley, 1980). Our recent companion papers have challenged this view, concluding that a signal combination strategy takes place across several grating cycles (Meese, 2010; Meese & Summers, 2007) and across eyes (Meese et al., 2006) at threshold and above. The study here has

extended the inquiry by combining these two dimensions to investigate the summation processes for suprathreshold signals presented to different eyes *and* different spatial (retinal) locations. Our results have confirmed that for sine-wave carrier gratings, spatial pooling is indifferent to eye of origin at threshold and above.

A three-stage model of contrast gain control fitted the results, where the contrast responses of individual filter elements were normalized before each of the signal summation stages. This gave the model the properties of ocularity invariance and area invariance above threshold, consistent with human perception, while leaving the benefits of signal summation intact at detection threshold. We also argue that this arrangement could form the basis of a population code for texture/pattern area.

An important step toward our conclusions and model development was the design of a novel stimulus class (the “Swiss cheese”; see also Meese & Summers, 2007, 2009, and the related “Battenberg” stimulus used by Meese, 2010) that overcame several design limitations that appear to have plagued previous studies. In particular, our approach suggests that the retinal regions of spatial integration and suppression were constant for the different stimulus conditions used here. One important direction for future work is to investigate the rules that control the reach of these pooling regions for more complex (e.g., natural) images.

## Appendix A

### Model details

The basic approach to the modeling here is the same as that described in Meese (2010) and Meese and Summers (2007, 2009).

Images had a contrast of 100% and were sampled with a resolution of 10 pixels per carrier cycle (though this was not critical) and multiplied by the attenuation surface shown in Figure 2b of Meese (2010) to simulate the effects of retinal inhomogeneity (Pointer & Hess, 1989).

The attenuated images were then filtered by Cartesian separable log-Gabor filters (Meese, 2010) with spatial frequency bandwidth of 1.6 octaves (full-width at half-height) and orientation bandwidth of  $\pm 25^\circ$  (half-width at half-height), though these parameters are not critical. The filters had phases from the set:  $\text{PHASE} = \{0^\circ, 90^\circ, 180^\circ, 270^\circ\}$ , where elements are denoted:  $\Phi$ . The filters were matched to the spatial frequency and orientation of the stimulus carriers for left and right eyes and were half-wave rectified.

With this formulation, the filter outputs were expressed in percentage units and represented the spatial distribution of filter-element (convolution kernel) responses across space for stimuli with contrasts of 100%. The responses for

specific contrast levels were then derived by multiplying this pattern of filter responses by the Michelson contrast (0:1) of the stimulus. This produced eight 2D arrays of contrast responses (in %) given by:  $L_{\text{PHASE}}, R_{\text{PHASE}}$ .

The Stage 1 output of each of the four phase filters for the left eye was derived as follows:

$$\text{Stage1}(\Phi, L, R) = L_{\Phi}^m / \left( S + \sum_{\phi \in \text{PHASE}} (L_{\phi}) + \sum_{\phi \in \text{PHASE}} (R_{\phi}) \right), \quad (\text{A1})$$

where the first nonlinear contrast transducer exponent  $m = 1.2$  (Meese & Summers, 2009). The pooling over phase on the denominator is consistent with the dichoptic masking results of Baker and Meese (2007) and is extended to include phase pooling from the ipsiocular channel, but this is not critical here.

An analogous operation was performed for the right eye. This was followed by binocular summation:

$$\text{Bin}(\Phi) = \text{Stage1}(\Phi, L, R) + \text{Stage1}(\Phi, R, L). \quad (\text{A2})$$

The second stage of gain control included summation across phase ( $\Phi$ ) and across  $n$  image locations ( $i$ ) (i.e., each of the pixels in the image):

$$\text{Stage2} = \frac{\sum_{i=1:n} \left[ \sum_{\Phi \in \text{PHASE}} (\text{Bin}(\Phi)_i^r) \right]}{B + \sum_{i=1:n} \left[ \sum_{\Phi \in \text{PHASE}} (\text{Bin}(\Phi)_i) \right]}, \quad (\text{A3})$$

where the second nonlinear contrast transducer exponent  $r = 2/m$ . The summation across area was critical for the results here; the summation across phase was not. Conceptually, we can envisage Equation A3 as follows. After binocular summation (Equation A2), each filter element is subject to nonlinear contrast transduction ( $r$ ) and is normalized by a large suppressive pool over area and phase. The responses of all these mechanisms are then summed.

In the simplest version of the model,  $B$  was a constant. This achieved good fits to the data here (RMS error of 1.0 dB) but did not achieve complete area invariance because even at the highest pedestal contrasts the denominator term did not completely overcome the influence of  $B$ . To solve this problem, we made  $B$  a function of the numerator term thus:

$$B = \frac{B'}{1 + \sum_{i=1:n} \left[ \sum_{\Phi \in \text{PHASE}} (\text{Bin}(\Phi)_i^r) \right]}. \quad (\text{A4})$$

Fixed parameters	Fitted parameters	Value
$S$		1.0
	$B'$	$117\pi\rho^2$
	$Z$	0.335
$m$		1.2
$r$		$2/m$
	$p$	12.35
	$q$	9.65
	$k$	0.0585

Table A1. Parameters for the three-stage model of contrast gain control. The radius of the stimulus ( $\rho$ ) was a constant and equal to 126 (pixels).

This improved the fit slightly (RMS error = 0.85 dB) and improved the area invariance of the model (Figure 7) without introducing any additional parameters.

The third and final stage of gain control is given by

$$\text{Stage3} = \frac{\text{Stage2}^p}{Z^q + \text{Stage2}^q}, \quad (\text{A5})$$

which is the output of the model and forms the decision variable. The model equations were solved for target contrast such that  $\text{Stage3}(\text{Pedestals} + \text{Targets}) - \text{Stage3}(\text{Pedestals}) = k$ , where  $k$  is a free parameter that sets the signal-to-noise ratio for detection and discrimination thresholds.

A full list of model parameter values is shown in Table A1.

## Appendix B

### The three-stage model of contrast gain control generalizes over other results

Here we compare the predictions of our three-stage model of contrast gain control with results from four different types of pedestal masking experiment (including this one). The parameter values are those used for the fitting in Figure 4. In the bottom row only, data and models have been slid within the axes to facilitate comparisons.

The three-stage model of contrast gain control provides a good account of the results from all four types of experiment in Figure B1. The main weakness is the overestimation in the level of area summation at low pedestal contrasts when grating area is manipulated (Figures B1g and B1h). This could be because the spatial extent of summation in the experiment was less than that in the model. However, another view anticipates this shortcoming from previous work where we have argued

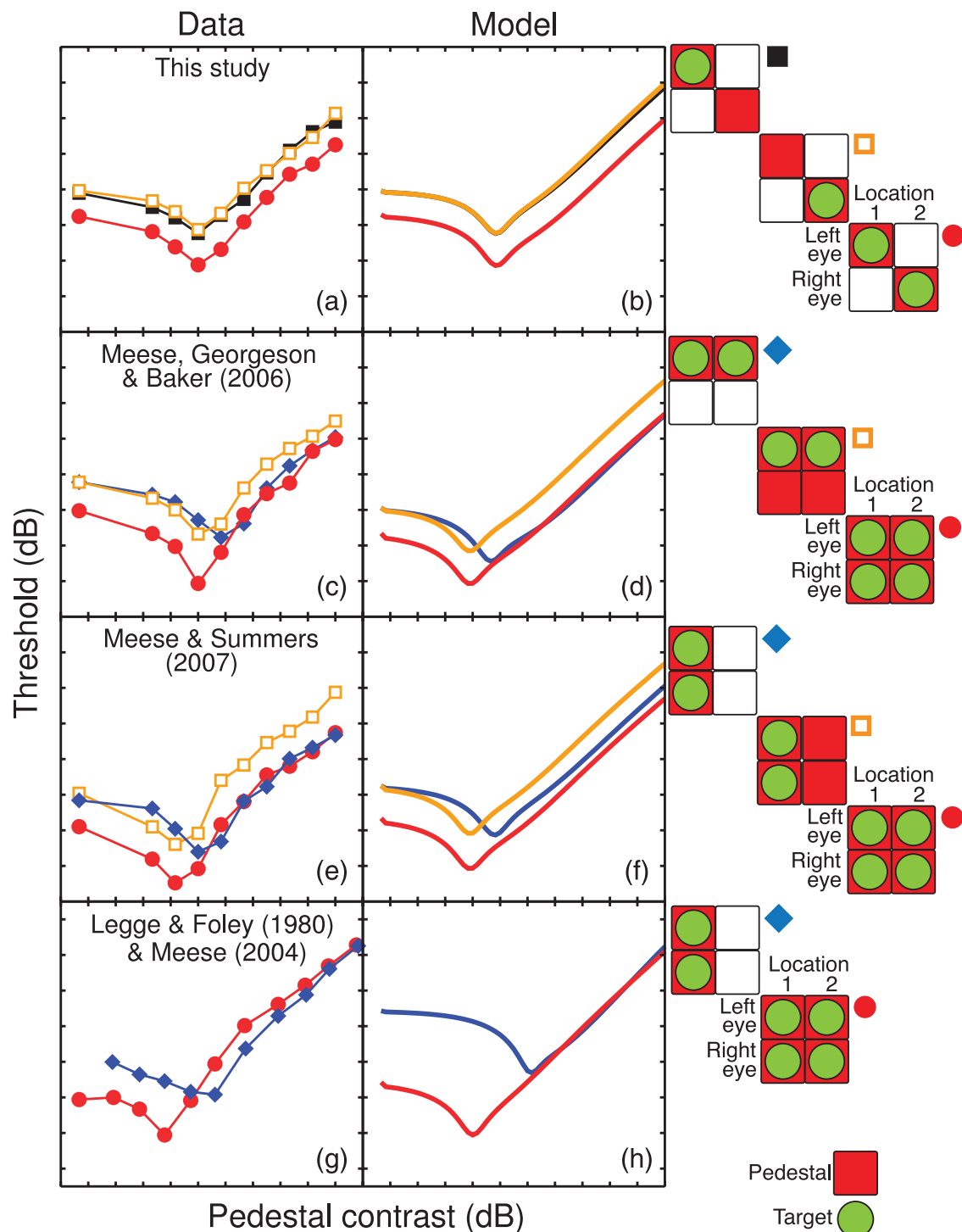


Figure B1. Psychophysical results and predictions from the three-stage model of contrast gain control for four different types of pedestal masking experiment. (a, b) Summation across eyes and area (using Swiss cheese stimuli). (c, d) Summation across eyes (using gratings). (e, f) Summation across area (using Swiss cheese stimuli). (g, h) Summation across area (using gratings). The data in (g) are replotted from Legge and Foley (1980). The stimuli used in the model were the circular patches of grating used in a similar study by Meese (2004) but with the areas used by Legge and Foley (1980). The stimulus conditions for each of the four rows are denoted by the icons on the right (see also Figure 1 in the main body of the report). Tick marks represent contrast steps of 6 dB (a factor of 2).



that summation of grating contrast at threshold is limited by the combined effects of nonlinear contrast transduction and integration of early noise. As mentioned in the main body of the report, this could have been addressed here by placing limiting additive noise just before the area summation stage in the model (Meese, 2010; Meese & Summers, 2007, 2009). This would reduce the level of model summation at threshold owing to the increase in noise for the large stimulus condition (Meese & Summers, 2007). However, the effects of subsequent model nonlinearities would then be lost to the performance measures owing to Birdsall's theorem (Klein & Levi, 2009; Lasley & Cohn, 1981). One plausible way out of this conundrum is that the compression of the contrast response—controlled here by the late (Stage 3) transducers—is actually controlled by late multiplicative noise, evident only above threshold.

## Acknowledgments

This work was supported by grants from the Engineering and Physical Sciences Research Council UK (GR/S74515/01; EP/H000038/1) and the Biotechnology and Biological Sciences Research Council UK (BB/H00159X/1), each awarded to Tim Meese and Mark Georgeson.

Commercial relationships: none.

Corresponding author: Dr. Tim S. Meese.

Email: t.s.meese@aston.ac.uk.

Address: Aston Triangle, Birmingham B4 7ET, UK.

## Footnotes

<sup>1</sup>There is an important difference between the models proposed by Foley et al. (2007) and Meese and Summers (2007). Meese and Summers proposed that noise propagates from each of the preliminary (V1-type) receptive fields, thereby increasing noise with summation area, whereas Foley et al. proposed that noise was constant with area. Both groups have presented evidence to support their proposals. Foley et al. found that thresholds for a small and a large grating patch were not affected by the choice of interleaved or blocked experimental design (see also Meese et al., 2005); if noise varies with stimulus size, the interleaved design should degrade performance for one or both of the two stimuli, depending upon attention strategy (see Tyler & Chen, 2000). However, in a preliminary report of a similar experiment with a greater number of stimulus patches (six), Summers and Meese (2007) found that performance was significantly worse for the smaller patches using an interleaved design compared to a blocked design. The reasons for these differences across studies are not clear. However, if intrinsic uncertainty in Foley et al.'s

experiment dominated the extrinsic uncertainty inherent in the interleaved design, then the design effect would be diluted and difficult to measure.

<sup>2</sup>By “contrast area,” we mean the spatial integral over the stimulus region of local contrast. This is distinct from contrast energy, which involves taking the square of local contrasts before integration.

<sup>3</sup>Our experiments involved the detection of first-order increments of luminance contrast both at and above threshold. However, because of the spatial modulation that we used (Equation 1), our stimuli also contained second-order components. These were discussed by Meese (2010) and Meese and Summers (2007), who concluded that they were probably of little or no importance for the type of experimental results here.

## References

- Abbey, C. K., & Eckstein, M. P. (2007). Classification images for simple detection and discrimination tasks in correlated noise. *Journal of the Optical Society of America A*, 24, B110–B124.
- Anzai, A., Peng, X., & Van Essen, D. C. (2007). Neurons in monkey visual area V2 encode combinations of orientations. *Nature Neuroscience*, 10, 1313–1321.
- Arcizet, F., Jouffrais, C., & Girard, P. (2008). Natural textures classification in area V4 of the macaque monkey. *Experimental Brain Research*, 189, 109–120.
- Baker, D. H., & Graf, E. W. (2009). On the relation between dichoptic masking and binocular rivalry. *Vision Research*, 49, 451–459.
- Baker, D. H., & Meese, T. S. (2007). Binocular contrast interactions: Dichoptic masking is not a single process. *Vision Research*, 47, 3096–3107.
- Baker, D. H., Meese, T. S., & Georgeson, M. A. (2007). Binocular interaction: Contrast matching and contrast discrimination are predicted by the same model. *Spatial Vision*, 20, 397–413.
- Baker, D. H., Meese, T. S., Mansouri, B., & Hess, R. F. (2007). Binocular summation of contrast remains intact in strabismic amblyopia. *Investigative Ophthalmology & Visual Science*, 48, 5332–5338.
- Baker, D. H., Meese, T. S., & Summers, R. J. (2007). Psychophysical evidence for two routes to suppression before binocular summation of signals in human vision. *Neuroscience*, 146, 435–448.
- Bergen, J. R., Wilson, H. R., & Cowen, J. D. (1979). Further evidence for four mechanisms mediating vision at threshold: Sensitivities to complex gratings and aperiodic stimuli. *Journal of the Optical Society of America*, 69, 1580–1587.

- Bonneh, Y., & Sagi, D. (1999). Contrast integration across space. *Vision Research*, 39, 2597–2602.
- Bowne, S. (1990). Contrast discrimination cannot explain spatial-frequency, orientation or temporal frequency discrimination. *Vision Research*, 30, 449–461.
- Cadieu, C., Kouh, M., Pasupathy, A., Conner, C. E., Riesenhuber, M., & Poggio, T. (2007). A model of V4 shape selectivity and invariance. *Journal of Neurophysiology*, 98, 1733–1750.
- Campbell, F. W., & Green, D. G. (1965). Monocular versus binocular visual acuity. *Nature*, 208, 191–192.
- Cannon, M. W., & Fullenkamp, S. C. (1988). Perceived contrast and stimulus size: Experiment and simulation. *Vision Research*, 28, 695–709.
- Cannon, M. W., & Fullenkamp, S. C. (1991a). A transducer model for contrast perception. *Vision Research*, 31, 983–998.
- Cannon, M. W., & Fullenkamp, S. C. (1991b). Spatial interactions in apparent contrast: Inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations. *Vision Research*, 31, 1985–1998.
- Cant, J. S., Arnott, S. R., & Goodale, M. A. (2009). fMR adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. *Experimental Brain Research*, 192, 391–405.
- Cavanaugh, J. R., Bair, W., & Movshon, A. (2002). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *Journal of Neurophysiology*, 88, 2530–2546.
- Chirimuuta, M., & Tolhurst, D. J. (2005). Does a Bayesian model of V1 contrast coding offer a neurophysiological account of human contrast discrimination? *Vision Research*, 45, 2943–2959.
- Cornsweet, T. N. (1962). The staircase method in psychophysics. *American Journal of Psychology*, 75, 485–491.
- Dakin, S. C., & Bex, P. J. (2002). Summation of concentric orientation structure: Seeing the glass or the window? *Vision Research*, 42, 2013–2020.
- Dakin, S. C., & Mareschal, I. (2000). Sensitivity to contrast modulation depends on carrier spatial frequency and orientation. *Vision Research*, 40, 311–329.
- Desimone, R., & Schein, S. J. (1987). Visual properties of neurons in area V4 of the macaque sensitivity to stimulus form. *Journal of Neurophysiology*, 57, 835–868.
- Dickinson, J. E., & Badcock, D. R. (2007). Selectivity for coherence in polar orientation in human form vision. *Vision Research*, 47, 3078–3087.
- Ding, J., & Sperling, G. (2006). A gain-control theory of binocular combination. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 1141–1146.
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual-system—Evidence for a local association field. *Vision Research*, 33, 173–193.
- Finn, I. M., & Ferster, D. (2007). Computational diversity in complex cells of cat primary visual cortex. *Journal of Neuroscience*, 27, 9638–9647.
- Finney, D. J. (1971). *Probit analysis* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A*, 11, 1710–1719.
- Foley, J. M., Varadharajan, S., Koh, C. C., & Farias, C. Q. (2007). Detection of Gabor patterns of different sizes, shapes, phases and eccentricities. *Vision Research*, 47, 85–107.
- Gallant, J. L., Conner, C. E., Rakshit, S., Lewis, J. W., & Van Essen, D. C. (1996). Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *Journal of Neurophysiology*, 76, 2718–2738.
- García-Pérez, M. A. (1988). Space-variant visual processing: Spatially limited visual channels. *Spatial Vision*, 3, 129–142.
- Georgeson, M. A., & Meese, T. S. (1997). Perception of stationary plaids: The role of spatial filters in edge analysis. *Vision Research*, 37, 3255–3271.
- Georgeson, M. A., & Meese, T. S. (1999). Perception of edges in plaids suggests adaptive filtering in human spatial vision. *Perception*, 28, 687–702.
- Georgeson, M. A., & Meese, T. S. (2005). Binocular summation at contrast threshold: A new look. *Perception*, 34, 138.
- Georgeson, M. A., & Meese, T. S. (2006). Fixed or variable noise in contrast discrimination? The jury's still out. *Vision Research*, 46, 4294–4303.
- Georgeson, M. A., & Meese, T. S. (2007). Binocular combination at threshold: Temporal filtering and summation of signals in separate ON and OFF channels. *Perception*, 36, 60.
- Georgeson, M. A., & Schofield, A. J. (2002). Shading and texture: Separate information channels with a common adaptation mechanism? *Spatial Vision*, 16, 59–76.
- Ghose, G. M., & Maunsell, J. H. R. (2008). Spatial summation can explain the Attentional modulation of neuronal responses to multiple stimuli in area V4. *Journal of Neuroscience*, 28, 5115–5126.

- Graham, N., & Sutter, A. (1998). Spatial summation in simple (Fourier) and complex (non-Fourier) texture channels. *Vision Research*, 38, 231–257.
- Grigorescu, C., Petkov, N., & Westenberg, M. A. (2004). Contour and boundary detection improved by surround suppression of texture edges. *Image and Vision Computing*, 22, 609–622.
- Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, 9, 181–197.
- Hess, R. F., & Field, D. J. (1995). Contour integration across depth. *Vision Research*, 35, 1699–1711.
- Huang, P. C., Hess, R. F., & Dakin, S. C. (2006). Flank facilitation and contour integration: Different sites. *Vision Research*, 46, 3699–3706.
- Itti, L., Koch, C., & Braun, J. (2000). Revisiting spatial vision: Toward a unifying model. *Journal of the Optical Society of America A*, 17, 1899–1917.
- Jones, D. G., Anderson, N. D., & Murphy, K. M. (2003). Orientation discrimination in visual noise using global and local stimuli. *Vision Research*, 43, 1123–1233.
- Kasamatsu, T., Miller, R., Zhu, Z., Chang, M., & Ishida, Y. (2010). Collinear facilitation is independent of receptive-field expansion at low contrast. *Experimental Brain Research*, 201, 453–465.
- Kawabe, T., & Miura, K. (2004). Configuration effects on texture transparency. *Spatial Vision*, 17, 187–200.
- Kersten, D. (1984). Spatial summation in visual noise. *Vision Research*, 24, 1977–1990.
- Kingdom, F. A. A., Prins, N., & Hayes, A. (2003). Mechanism independence for texture-modulation detection is consistent with a filter–rectify–filter mechanism. *Visual Neuroscience*, 20, 65–76.
- Klein, S. A., & Levi, D. M. (2009). Stochastic model for detection of signals in noise. *Journal of the Optical Society of America A*, 26, B110–B126.
- Kontsevich, L. L., Chen, C.-C., & Tyler, C. W. (2002). Separating the effects of response nonlinearity and internal noise psychophysically. *Vision Research*, 42, 1771–1784.
- Kontsevich, L. L., & Tyler, C. W. (1994). Analysis of stereothresholds for stimuli below 2.5 c/deg. *Vision Research*, 34, 2317–2329.
- Köteles, K., De Mazière, P. A., Van Hulle, M., Orban, G. A., & Vogels, R. (2008). Coding of images of materials by macaque inferior temporal cortical neurons. *European Journal of Neuroscience*, 27, 466–482.
- Kouh, M., & Poggio, T. (2008). A canonical neural circuit for cortical nonlinear operations. *Neural Computation*, 20, 1427–1451.
- Lampl, I., Ferster, D., Poggio, T., & Riesenhuber, M. (2004). Intracellular measurements of spatial integration and the MAX operation in complex cells of the cat primary visual cortex. *Journal of Neurophysiology*, 92, 2704–2713.
- Lasley, D. J., & Cohn, T. E. (1981). Why luminance discrimination may be better than detection. *Vision Research*, 21, 273–278.
- Legge, G., & Foley, J. (1980). Contrast masking in human vision. *Journal of the Optical Society of America A*, 70, 1458–1471.
- Legge, G. E. (1979). Spatial-frequency masking in human vision—Binocular interactions. *Journal of the Optical Society of America*, 69, 838–847.
- Legge, G. E. (1984). Binocular contrast summation—II. Quadratic summation. *Vision Research*, 24, 385–394.
- Legge, G. E., & Rubin, G. S. (1981). Binocular interactions in suprathreshold contrast perception. *Perception & Psychophysics*, 30, 49–61.
- Levi, D. M., & Klein, S. A. (2000). Seeing circles: What limits shape perception? *Vision Research*, 40, 2329–2339.
- Levi, D. M., Klein, S. A., & Chen, C.-C. (2005). What is the signal in noise? *Vision Research*, 45, 1835–1846.
- Lu, Z. L., & Doshier, B. A. (2008). Characterizing observers using external noise and observer models: Assessing internal representations with external noise. *Psychological Review*, 115, 44–81.
- Luntinen, O., Rovamo, J., & Näsänen, R. (1995). Modelling the increase of contrast sensitivity with grating area and exposure time. *Vision Research*, 35, 2339–2346.
- Maehara, G., & Goryo, K. (2005). Binocular, monocular and dichoptic pattern masking. *Optical Review*, 12, 76–82.
- Manahilov, V., Simpson, W. A., & McCulloch, D. L. (2001). Spatial summation of peripheral Gabor patches. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, 18, 273–282.
- Mansouri, B., Hess, R. F., Allen, H. A., & Dakin, S. C. (2005). Integration, segregation, and binocular combination. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, 22, 38–48.
- Mareschal, I., & Baker, C. L. (1998). A cortical locus for the processing of contrast-defined contours. *Nature Neuroscience*, 1, 150–154.
- Mareschal, I., Morgan, M. J., & Solomon, J. A. (2010). Cortical distance determines whether flankers cause crowding or the tilt illusion. *Journal of Vision*, 10(8):13, 1–14. <http://www.journalofvision.org/content/10/8/13>, doi:10.1167/10.8.13. [PubMed] [Article]



- May, K. A., & Hess, R. F. (2007a). Dynamics of snakes and ladders. *Journal of Vision*, 7(12):13, 1–9, <http://www.journalofvision.org/content/7/12/13>, doi:10.1167/7.12.13. [PubMed] [Article]
- May, K. A., & Hess, R. F. (2007b). Ladder contours are undetectable in the periphery: A crowding effect? *Journal of Vision*, 7(13):9, 1–15, <http://www.journalofvision.org/content/7/13/9>, doi:10.1167/7.13.9. [PubMed] [Article]
- Mayer, M. J., & Tyler, C. W. (1986). Invariance of the slope of the psychometric function with spatial summation. *Journal of the Optical Society of America A*, 3, 1166–1172.
- McIlhagga, W., & Pääkkönen, A. (1999). Noisy templates explain area summation. *Vision Research*, 39, 367–372.
- Meese, T. S. (2004). Area summation and masking. *Journal of Vision*, 4(10):8, 930–943, <http://www.journalofvision.org/content/4/10/8>, doi:10.1167/4.10.8. [PubMed] [Article]
- Meese, T. S. (2010). Spatially extensive summation of contrast energy is revealed by contrast detection of micro-pattern textures. *Journal of Vision*, 10(8):14, 1–21, <http://www.journalofvision.org/content/10/8/14>, doi:10.1167/10.8.14. [PubMed] [Article]
- Meese, T. S., & Baker, D. H. (2009). Cross-orientation masking is speed invariant between ocular pathways but speed dependent within them. *Journal of Vision*, 9(5):2, 1–15, <http://www.journalofvision.org/content/9/5/2>, doi:10.1167/9.5.2. [PubMed] [Article]
- Meese, T. S., Challinor, K. L., & Summers, R. J. (2008). A common contrast pooling rule for suppression within and between the eyes. *Visual Neuroscience*, 25, 585–601.
- Meese, T. S., & Georgeson, M. A. (2005). Carving up the patchwise transform: Towards a filter combination model for spatial vision. *Advances in Psychology Research*, 34, 51–88.
- Meese, T. S., Georgeson, M. A., & Baker, D. H. (2006). Binocular contrast vision at and above threshold. *Journal of Vision*, 6(11):7, 1224–1243, <http://www.journalofvision.org/content/6/11/7>, doi:10.1167/6.11.7. [PubMed] [Article]
- Meese, T. S., & Hess, R. F. (2007). Anisotropy for spatial summation of elongated patches of grating: A tale of two tails. *Vision Research*, 47, 1880–1892.
- Meese, T. S., Hess, R. F., & Williams, B. W. (2005). Size matters, but not for everyone: Individual differences for contrast discrimination. *Journal of Vision*, 5(11):2, 928–947, <http://www.journalofvision.org/content/5/11/2>, doi:10.1167/5.11.2. [PubMed] [Article]
- Meese, T. S., & Holmes, D. J. (2004). Performance data indicate summation for pictorial depth cues in slanted surfaces. *Spatial Vision*, 17, 127–151.
- Meese, T. S., & Summers, R. J. (2007). Area summation in human vision at and above detection threshold. *Proceedings of the Royal Society B*, 274, 2891–2900.
- Meese, T. S., & Summers, R. J. (2009). Neuronal convergence in early contrast vision: Binocular summation is followed by response nonlinearity and area summation. *Journal of Vision*, 9(4):7, 1–16, <http://www.journalofvision.org/content/9/4/7>, doi:10.1167/9.4.7. [PubMed] [Article]
- Meese, T. S., & Williams, C. B. (2000). Probability summation for multiple patches of luminance modulation. *Vision Research*, 40, 2101–2113.
- Moradi, F., & Heeger, D. J. (2009). Inter-ocular contrast normalization in human visual cortex. *Journal of Vision*, 9(3):13, 1–22, <http://www.journalofvision.org/content/9/3/13>, doi:10.1167/9.3.13. [PubMed] [Article]
- Morrone, M. C., Burr, D. C., & Vaina, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature*, 376, 507–509.
- Motoyoshi, I., & Kingdom, F. A. A. (2007). Differential roles of contrast polarity reveal two streams of second-order visual processing. *Vision Research*, 47, 2047–2054.
- Motoyoshi, I., & Nishida, S. Y. (2004). Cross-orientation summation in texture segregation. *Vision Research*, 44, 2567–2576.
- Nachmias, J. (2002). Contrast discrimination with and without spatial uncertainty. *Vision Research*, 42, 41–48.
- Näsänen, R., Tiippana, K., & Rovamo, J. (1998). Contrast restoration model for contrast matching of cosine gratings of various spatial frequencies and areas. *Ophthalmic and Physiological Optics*, 18, 269–278.
- Nurminen, L., Kilpeläinen, M., Laurinen, P., & Vanni, S. (2009). Area summation in human visual system: Psychophysics, fMRI and modeling. *Journal of Neurophysiology*, 102, 2900–2909.
- Ostwald, D., Lam, J. M., Li, S., & Kourtzi, Z. (2008). Neural coding of global form in human visual cortex. *Journal of Neurophysiology*, 99, 2456–2469.
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, 4, 739–744.
- Pearson, J., Tadin, D., & Blake, R. (2007). The effects of transcranial magnetic stimulation on visual rivalry. *Journal of Vision*, 7(7):2, 1–11, <http://www.journalofvision.org/content/7/7/2>, doi:10.1167/7.7.2. [PubMed] [Article]
- Peirce, J. W., Solomon, S. G., Forte, J. D., & Lennie, P. (2008). Cortical representation of color is binocular. *Journal of Vision*, 8(3):6, 1–10, <http://www.journalofvision.org/content/8/3/6>, doi:10.1167/8.3.6. [PubMed] [Article]

- Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A*, 2, 1508–1532.
- Petrov, A. A., Dosher, B. A., & Lu, Z. L. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review*, 112, 715–743.
- Pointer, J. S., & Hess, R. F. (1989). The contrast sensitivity gradient across the human visual field—With emphasis on the low spatial-frequency range. *Vision Research*, 29, 1133–1151.
- Pollen, D. A., Przybyszewski, A. W., Rubin, M. A., & Foote, W. (2002). Spatial receptive field organization of macaque V4 neurons. *Cerebral Cortex*, 12, 601–616.
- Press, W. A., Brewer, A. A., Dougherty, R. F., Wade, A. R., & Wandell, B. A. (2001). Visual areas and spatial summation in human visual cortex. *Vision Research*, 41, 1321–1332.
- Priebe, N. J., & Ferster, D. (2008). Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. *Neuron*, 57, 482–497.
- Quick, R. F. (1974). A vector-magnitude model of contrast detection. *Kybernetik*, 16, 65–67.
- Read, J. C. A., Parker, A. J., & Cumming, B. G. (2002). A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images. *Visual Neuroscience*, 19, 735–753.
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2, 1019–1025.
- Riesenhuber, M., & Poggio, T. (2002). Neural mechanisms of object recognition. *Current Opinion in Neurobiology*, 12, 162–168.
- Ringach, D. L. (2010). Population coding under normalization. *Vision Research*, 50, 2223–2232.
- Roach, N. W., Webb, B. S., & McGraw, P. V. (2008). Adaptation to global structure induces spatially remote distortions of perceived orientation. *Journal of Vision*, 8(3):31, 1–12, <http://www.journalofvision.org/content/8/3/31>, doi:10.1167/8.3.31. [PubMed] [Article]
- Robson, J. G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research*, 21, 409–418.
- Rose, D. (1980). The binocular/monocular sensitivity ratio for movement direction varies with temporal frequency. *Perception*, 9, 577–580.
- Rovamo, J., Luntinen, O., & Nasanen, R. (1993). Modelling the dependence of contrast sensitivity on grating area and spatial frequency. *Vision Research*, 33, 2773–2788.
- Rovamo, J., Mustonen, J., & Nasanen, R. (1994). Modeling contrast sensitivity as a function of retinal illuminance and grating area. *Vision Research*, 34, 1301–1314.
- Rovamo, J., Ukkonen, O., Thompson, C., & Nasanen, R. (1994). Spatial integration of compound gratings with various numbers of orientation components. *Investigative Ophthalmology & Visual Science*, 35, 2611–2619.
- Saarinen, J., Levi, D. M., & Shen, B. (1997). Integration of local pattern elements into global shape in human vision. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 8267–8271.
- Sachs, M. B., Nachmias, J., & Robson, J. G. (1971). Spatial-frequency channels in human vision. *Journal of the optical Society of America*, 61, 1176–1186.
- Sayim, B., Westheimer, G., & Herzog, M. H. (2010). Gestalt factors modulate basic spatial vision. *Psychological Science*, 21, 641–644.
- Sclar, G., Maunsell, J. H. R., & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research*, 30, 1–10.
- Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature*, 378, 492–496.
- Simmons, D. R. (2005). The binocular combination of chromatic contrast. *Perception*, 34, 1035–1042.
- Solomon, J. A. (2002). Noise reveals visual mechanisms of detection and discrimination. *Journal of Vision*, 2(1):7, 105–120, <http://www.journalofvision.org/content/2/1/7>, doi:10.1167/2.1.7. [PubMed] [Article]
- Summers, R. J., & Meese, T. S. (2006). Summation for pictorial and non-pictorial routes to depth perception. *Perception*, 35, 427.
- Summers, R. J., & Meese, T. S. (2007). Area summation is linear but the contrast transducer is nonlinear: Models of summation and uncertainty and evidence from the psychometric function. *Perception*, 36, 5.
- Swanson, W. H., Wilson, H. R., & Giese, S. C. (1984). Contrast matching data predicted from contrast increment thresholds. *Vision Research*, 24, 63–75.
- Tolhurst, D. J. (2007). Trying to model grating-contrast discrimination dippers and natural-scene discriminations. *Perception*, 36, 311–312.
- Tyler, C. W., & Chen, C. C. (2000). Signal detection theory in the 2AFC paradigm: Attention, channel uncertainty and probability summation. *Vision Research*, 40, 3121–3144.
- Van Boxtel, J. J. A., Van Ee, R., & Erkelens, C. J. (2007). Dichoptic masking and binocular rivalry share common perceptual dynamics. *Journal of Vision*, 7(14):3, 1–11, <http://www.journalofvision.org/content/7/14/3>, doi:10.1167/7.14.3. [PubMed] [Article]
- Watson, A. B. (1979). Probability summation over time. *Vision Research*, 19, 515–522.
- Watson, A. B., & Ahumada, A. J. (2005). A standard model for foveal detection of spatial contrast. *Journal*

- of Vision, 5, 717–740, <http://www.journalofvision.org/content/5/9/6>, doi:10.1167/5.9.6. [PubMed] [Article]
- Watson, A. B., & Robson, J. G. (1981). Discrimination at threshold: Labelled detectors in human vision. *Vision Research*, 21, 1115–1122.
- Watt, R. J., & Morgan, M. J. (1984). Spatial filters and the localization of luminance changes in human vision. *Vision Research*, 24, 1387–1397.
- Webb, B. S., Roach, N. W., & Peirce, J. W. (2008). Masking exposes multiple global form mechanisms. *Journal of Vision*, 8(9):16, 1–10, <http://www.journalofvision.org/content/8/9/16>, doi:10.1167/8.9.16. [PubMed] [Article]
- Wetherill, G. B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *Journal of Experimental Psychology*, 15, 485–492.
- Wilkinson, F., Wilson, H. R., & Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America A*, 14, 2057–2068.
- Wilkinson, F., Wilson, H. R., & Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research*, 38, 3555–3568.
- Williams, A. L., Singh, K. D., & Smith, A. T. (2003). Surround modulation measured with functional MRI in the human visual cortex. *Journal of Neurophysiology*, 89, 525–533.
- Wilson, H. R. (2003). Computational evidence for a rivalry hierarchy in vision. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 14499–14503.
- Wilson, H. R., & Bergen, J. R. (1979). A mechanism model for threshold spatial vision. *Vision Research*, 19, 19–32.
- Wilson, H. R., McFarlane, D. K., & Phillips, G. C. (1983). Spatial-frequency tuning of orientation selective units estimated by oblique masking. *Vision Research*, 23, 873–882.
- Wilson, H. R., & Wilkinson, F. (1998). Detection of global structure in Glass patterns: Implications for form vision. *Vision Research*, 38, 2933–2947.
- Wilson, H. R., Wilkinson, F., & Asaad, W. (1997). Concentric orientation summation in human form vision. *Vision Research*, 37, 2325–2330.
- Xing, J., & Heeger, D. J. (2000). Centre-surround interactions in foveal and peripheral vision. *Vision Research*, 40, 3065–3072.